

KIERUNKI ROZWOJU INSTRUMENTALNEJ I NIEINSTRUMENTALNEJ DETEKCJI KŁAMSTWA

**problemy kryminalistyczne,
etyczne i prawne**

KIERUNKI ROZWOJU INSTRUMENTALNEJ I NIEINSTRUMENTALNEJ DETEKCJI KŁAMSTWA

**problemy kryminalistyczne,
etyczne i prawne**

**redakcja
Jan Widacki**

**Autorzy:
Jacek Antos, Karolina Dukała, Marcin Gołaszewski, Aleksandra Kuhny,
Paweł Miazga, Zbigniew Mikrut, Romuald Polczyk, Anna Szuba-Boroń,
Marcin Tarabuła, Jan Widacki, Michał Widacki**

Rada Wydawnicza Krakowskiej Akademii im. Andrzeja Frycza Modrzewskiego:
Klemens Budzowski, Maria Kapiszewska, Zbigniew Maciąg, Jacek M. Majchrowski

Publikacja powstała w ramach projektu sfinansowanego ze środków Narodowego Centrum Nauki
przyznanych na podstawie decyzji numer DEC-2013/11/B/HS5/03856.

Recenzja: dr hab. Marek Leśniak

Projekt okładki: Oleg Aleksejczuk

ISBN 978-83-65208-91-0

Copyright© by Krakowska Akademia im. Andrzeja Frycza Modrzewskiego
Kraków 2018

Żadna część tej publikacji nie może być powielana ani magazynowana
w sposób umożliwiający ponowne wykorzystanie,
ani też rozpowszechniana w jakiegokolwiek formie
za pomocą środków elektronicznych, mechanicznych, kopiujących,
nagrywających i innych, bez uprzedniej pisemnej zgody właściciela praw autorskich.

Na zlecenie:



Krakowskiej Akademii
im. Andrzeja Frycza Modrzewskiego
www.ka.edu.pl

Wydawca: Oficyna Wydawnicza AFM, Kraków 2018

Sprzedaż prowadzi:
e-mail: ksiegarnia@kte.pl

Druk i oprawa: MKpromo

Spis treści

WPROWADZENIE	7
---------------------------	----------

Jan Widacki, Karolina Dukata

CZĘŚĆ I. WSTĘPNE USTALENIA TERMINOLOGICZNE.....	11
--	-----------

ANEKS A

– <i>Jan Widacki, Karolina Dukata</i> : Detekcja kłamstwa – czyli czego?	13
– <i>Jan Widacki, Karolina Dukata</i> : Lie detection – of what?	31
– <i>Jan Widacki</i> : Poligraf czy wariograf?	49
– <i>Jan Widacki</i> : Polygraph or Variograph?	55
– <i>Jan Widacki</i> : The First Polygraph?	59

*Jan Widacki, Romuald Polczyk, Marcin Gołaszewski, Michał Widacki, Paweł Miazga,
Karolina Dukata, Anna Szuba-Boroń, Marcin Tarabula, Aleksandra Kuhny*

CZĘŚĆ II. EKSPERYMENT	63
------------------------------------	-----------

1. Cel eksperymentu	63
2. Osoby badane z podziałem na role	65
3. Eksperci-poligraferzy	66
4. Aparat	67
5. Pokój do badań	67
6. Technika badania	68
6.1. Geneza techniki Utah	68
6.2. Główne założenia i struktura testu Utah ZCT	69
6.3. Metodyka przeprowadzania testu Utah ZCT	70
6.4. Ewaluacja danych testowych	72
6.5. Wartość diagnostyczna i dowodowa	74
7. Wyniki badań poligraficznych	74

ANEKS B

– <i>Marcin Tarabula, Michał Widacki</i> : The Amount of Information Remembered by the Perpetrator in the Context of the Application of the Guilty Knowledge Technique in Criminal Investigation – a Pilot Study	77
– <i>Jan Widacki</i> : Spór o technikę badań poligraficznych	87

Jan Widacki, Marcin Gołaszewski, Karolina Dukata, Romuald Polczyk

CZĘŚĆ III. SUBIEKTYWIZM W BADANIACH POLIGRAFICZNYCH	97
--	-----------

ANEKS C

– <i>Jan Widacki, Marcin Gołaszewski</i> : Subiektywizm w badaniach poligraficznych	99
---	----

Karolina Dukala, Jan Widacki

CZĘŚĆ IV. NIEINSTRUMENTALNE METODY DETEKCJI KŁAMSTWA	117
1. Wprowadzenie.....	117
2. Aktualny stan wiedzy	117
2.1. Obserwacja zachowania niewerbalnego	119
2.2. Werbalne narzędzia oceniania szczerości zeznań.....	121
2.3. Obciążenie poznawcze (<i>Cognitive Load Approach</i>).....	126
3. Przyszłość.....	127
4. Kwestie prawne związane z możliwością wykorzystania nieinstrumentalnych metod detekcji kłamstwa	128

Jan Widacki, Zbigniew Mikrut, Michał Widacki, Jacek Antos

CZĘŚĆ V. PRÓBA WYKORZYSTANIA ZMIAN TEMPERATURY TWARZY JAKO WSKAŹNIKA ZMIAN EMOCJONALNYCH PRZY DETEKCJI KŁAMSTWA	131
1. Wprowadzenie	131
2. Eliminacja wpływu niewielkich ruchów twarzy osoby badanej	134
3. Rejestracja poligraficzna a wyniki detekcji termowizyjnej	136
4. Wpływ niewielkich zmian położenia i rozmiarów okien analizy dla jakości zapisu	139
5. Analiza rejestracji (zapisów) uzyskanych podczas wykonywania testu Utah ZCT	143
6. Podsumowanie	147
– <i>Zbigniew Mikrut, Michał Widacki, Jan Widacki</i> Próba wykorzystania zmiany temperatury twarzy jako wskaźnika w instrumentalnej detekcji kłamstwa.....	133

ANEKS D

– <i>Marcin Gołaszewski, Paweł Zajac, Jan Widacki: Thermal Vision as a Method of Detection of Deception: A Review of Experiences.....</i>	149
– <i>Jan Widacki, Michał Widacki, Jacek Antos: Preparation to Experimental Testing of the Potential from Using Facial Temperature Changes Registered with an Infrared Camera in Lie Detection</i>	163

Jan Widacki

CZĘŚĆ VI. ASPEKTY PRAWNE I ETYCZNE	169
1. Wprowadzenie	169
2. Badania poligraficzne metodą bezkontaktową. Problemy techniczne i prawne	169

ANEKS E

– <i>Jan Widacki: Tezy postanowienia Sądu Najwyższego z dnia 29 stycznia 2015 r., I KZP 25/14.....</i>	177
– <i>Jan Widacki, Anna Szuba-Boroń: Badania poligraficzne w procesie karnym w świetle postanowienia Sądu Najwyższego z dnia 29 stycznia 2015 (I KZP 25/14)</i>	183
– <i>Jan Widacki: Opinia z badań poligraficznych w procesie karnym</i>	191
– <i>Anna Szuba-Boroń, Marcin Gołaszewski: Badania poligraficzne funkcjonariuszy państwowych w Polsce.....</i>	197

PODSUMOWANIE	203
---------------------------	------------

Wprowadzenie

Detekcja kłamstwa ma długą historię. Pierwsze próby detekcji kłamstwa w oparciu o osiągnięcia nauki, głównie psychologii i fizjologii, podejmowano już na przełomie XIX i XX wieku¹. Z czasem wykształciły się jej dwie podstawowe metody. Pierwsza z nich opierała się na obserwacji zachowania osoby, której wypowiedź oceniano, analizie jej zachowania, w szczególności ruchów wyrazowych (mimicznych, pantomimicznych), analizie sposobu mówienia, analizie treści jej wypowiedzi. Metoda ta nie wymagała używania żadnych instrumentów, żadnej aparatury. Druga polegała na obserwacji (później także rejestracji) fizjologicznych korelatów emocji osoby badanej, towarzyszących jej wypowiedziom. Wymagało to użycia specjalnej aparatury. Jednym z pierwszych takich aparatów był skonstruowany przez Angelo Mosso hydropletysmograf², pozwalający na obserwację zmian objętości przedramienia osoby badanej, będących wynikiem zmian ukrwienia, wywołanych emocjami, które towarzyszyły odpowiadaniu na pytania. Sfigmograf pozwalał na obserwację zmian częstotliwości tętna i ciśnienia tętniczego krwi, a pneumograf – obserwację zmian w przebiegu czynności oddychania (częstotliwość cykli oddechowych, zmiany głębokości oddechu). Połączenie tych urządzeń z kimografem pozwalało na rejestrację obserwowanych zmian³. Od lat 20. XX wieku, początkowo w Stanach Zjednoczonych, a po II wojnie światowej w całym niemal świecie rutynowo stosowany jest poligraf, urządzenie łączące sfigmograf i pneumograf, nieco później uzupełnione psychogalwanometrem⁴. Jest rzeczą ciekawą, że obydwie metody, instrumentalna i nieinstrumentalna, rozwijały się niezależnie od siebie, choć były podejmowane próby równoczesnego ich stosowania⁵.

W instrumentalnej detekcji kłamstwa na poziomie psychofizjologicznym próbowano wykorzystywać obserwacje innych fizjologicznych korelatów emocji niż te, które są przedmiotem obserwacji i rejestracji przy klasycznym badaniu poligraficznym. Polegały one na obserwacji zmian napięcia mięśni i drżenia ciała, czy obserwacji zmian brzmienia głosu. Od końca lat 70. postulowano możliwość wykorzystania kamery termowizyjnej do obserwacji i rejestracji zmian temperatury twarzy jako kolejnego fizjologicznego korelatu emocji, przydatnej do detekcji kłamstwa. Te ostatnie wydawały się szczególnie interesujące, mogłyby być wykorzystane nie tylko przy detekcji kłamstwa, ale także do zdalnej obserwacji twarzy ludzi uczestniczących w zgromadzeniach, przebywających w terminalach lotniczych i podobnych miejscach publicznych, dla typowania osób, które z jakichś powodów są podenerwowane, przeżywają podwyższone napięcie emocjonalne itp.

¹ Por. J. Widacki, *Historia badań poligraficznych*, Oficyna Wydawnicza AFM, Kraków 2017, s. 65 i nast.

² Por. *ibidem*, s. 68.

³ Por. *ibidem*, s. 42.

⁴ Por. *ibidem*, s. 84 i nast.

⁵ Por. np. F. Horvath, *Verbal and nonverbal clues to truth and deception during polygraph examination*, „Journal of Police Science and Administration” 1973, 1, 2, s. 138–152.

Możliwość wykorzystania do takich celów kamery termowizyjnej jest sygnalizowana w literaturze. W Polsce nie mieliśmy żadnych doświadczeń w tym zakresie, co utrudniało nam nawet krytyczne podejście do zagranicznych doniesień naukowych dotyczących tego zagadnienia.

Współcześnie w świecie prowadzone są badania naukowe zarówno w zakresie nieinstrumentalnej, jak i instrumentalnej detekcji kłamstwa.

W zakresie tej pierwszej wypracowano szereg metod, takich jak BAI (Behavioral Anaysis Interview), SVA/CBCA (Statement Validity Assessment/Criteria Based Content) czy SCAN (Scientific Content Analysis)⁶. Metody te są ciągle doskonalone. Równocześnie dyskutowane jest, w jakiej formie mogą być one wprowadzone w polskim postępowaniu karnym – jako metody pomocnicze dla przesłuchującego czy raczej jako narzędzie pracy biegłego.

W zakresie doskonalenia metod instrumentalnych można wyróżnić trzy podstawowe kierunki. Pierwszy z nich to doskonalenie tradycyjnego badania poligraficznego, polegające na eliminowaniu celowych zakłóceń zapisu wprowadzanych przez badanego, poprzez stosowanie coraz to lepszych, czulszych metod kontrolnego sprawdzania napięcia jego mięśni, na ulepszaniu technik badania i metod numerycznej oceny zapisów. Istotnym problemem jest też sprawdzenie zakresu subiektywizmu eksperta oceniającego zapis badania poligraficznego. Ma to znaczenie zarówno przy ocenie wartości diagnostycznej takiego badania, jak i przy określeniu możliwości skontrolowania zapisu z badania przez innego eksperta i wartości tej kontroli, a także dla ustalenia, na ile badanie poligraficzne dostarcza obiektywnego dowodu naukowego.

Drugi kierunek to szukanie symptomów kłamstwa nie na poziomie psychofizjologicznym, ale na poziomie neurofizjologicznym. Wraca się tu zarówno do zarzuconych swego czasu badań EEG, jak i próbuje się wykorzystać w celu detekcji kłamstwa obserwacje pracy mózgu za pomocą funkcjonalnego magnetycznego rezonansu jądrowego (functional magnetic resonance imaging, fMRI).

Kierunek trzeci obejmuje próby zastąpienia wybranych trzech fizjologicznych korelatów emocji (przebieg czynności oddychania, przebieg pracy serca, zmiany w zakresie odruchu skórno-galwanicznego) trzema innymi, których obserwacja i rejestracja nie wymaga zakładania na ciele badanego żadnych czujników, a zatem teoretycznie pozwala na wykonanie badania poligraficznego bez wiedzy, a tym samym zgody badanego. Najczęściej wymieniane są w tym kontekście zmiany w brzmieniu głosu, zmiany temperatury twarzy, ruchy gałek ocznych i odruch źreniczny.

Nie trzeba dodawać, że takie badanie, które mogłoby być wykonane bez wiedzy (i zgody!) badanego, wymagałoby innej od dotąd obowiązującej reglamentacji prawnej, wszak obecnie wykonanie badania poligraficznego bez wiedzy i zgody badanego jest faktycznie niemożliwe, a prawnie niedopuszczalne.

Z założenia grant NCN DEC-2013/11/B/HS5/03856 „**Kierunki rozwoju instrumentalnej i nieinstrumentalnej detekcji kłamstwa – problemy kryminalistyczne, etyczne i prawne**” nie miał przynieść gotowych rozwiązań przedstawionych kwestii, ale przybliżyć nas do nauki światowej na tyle, by możliwa była z nią kompetentna dyskusja, jej osiągnięcia w tym zakresie mogły być zrozumiałe, a rozważania o koniecznych rozwiązaniach prawnych dotyczyły realnych problemów, wynikających ze stanu badań. Realizacja projektu badawczego wymagała dokładnego przeanalizowania stanu nauki światowej, dokonania niezbędnych uściśleń terminologicznych i przeprowadzenia badań empirycznych, w tym eksperymentalnych.

⁶ Por. Kryminalistyka, wyd. 3, red. J. Widacki, C.H. Beck, Warszawa 2016, s. 144–146.

Wykorzystanie kamery termowizyjnej w celu obserwacji i rejestracji zmian temperatury twarzy dokonujących się pod wpływem emocji okazało się trudniejsze, niż początkowo zakładaliśmy. Podstawowe problemy udało nam się pokonać, a prowadzone badania wzbudziły zainteresowanie także za granicą. Jednak od momentu, do którego doszliśmy, do wykorzystania kamery jako instrumentu użytecznego przy praktycznej detekcji kłamstwa droga jeszcze daleka. Wymaga to dalszych badań eksperymentalnych i dalszego doskonalenia metody.

Niektóre ustalenia i częściowe wyniki badań publikowaliśmy na bieżąco. Zawierające je prace zamieszczone zostały w aneksach. Stanowią one konieczne i logiczne uzupełnienie narracji tej monografii.

Część I

Wstępne ustalenia terminologiczne

W tytule zadania badawczego użyto terminów „instrumentalna detekcja kłamstwa” oraz „nieinstrumentalna detekcja kłamstwa”. Obydwa wymagają wyjaśnienia i sprecyzowania. Zacząć trzeba od pojęcia „detekcja kłamstwa”. Oczywiście, przy aktualnym stanie nauki detekcja, czyli wykrywanie „kłamstwa” (*lie, deception*) jako takiego nie jest możliwa. Wykrywa się i analizuje zjawiska, które kłamstwu towarzyszą (napiecie emocjonalne, wysiłek intelektualny etc.) i manifestują się w różny sposób. Z wystąpienia tych towarzyszących kłamstwu zjawisk wnioskujemy o kłamstwie.

$$\{[(l \rightarrow a) \wedge (a \rightarrow l)] \wedge a\} \rightarrow l$$

gdzie:

a – symptomy towarzyszące kłamstwu

l – kłamstwo.

Sam termin „kłamstwo” też wymaga objaśnienia i doprecyzowania znaczenia, w jakim będziemy go używać. Następnie należy wyjaśnić, co w poszczególnych badaniach (instrumentalnych, nieinstrumentalnych, a nawet w różnych technikach tych badań, jest uznawane za „symptom towarzyszący kłamstwu”.

Jak już wiemy, Instrumentalna detekcja kłamstwa wymaga użycia aparatury. Metody nieinstrumentalnej detekcji kłamstwa nie posługują się z zasady żadną aparaturą. Polegają one na obserwacji symptomów werbalnych bądź niewerbalnych (behawioralnych) i na analizach wypowiedzi (ustnych lub pisemnych).

Instrumentalna detekcja kłamstwa, jak wspomniano we Wstępie, obejmuje przede wszystkim takie metody jak badania poligraficzne (rozwinęta i stosowana rutynowo metoda detekcji kłamstwa), pozwalające na obserwację i rejestrację tzw. fizjologicznych korelatów emocji, czyli zjawisk fizjologicznych towarzyszących zmianom emocjonalnym i ewentualnemu wysiłkowi intelektualnemu, które współwystępują z kłamstwem w specyficznych, wyreżyserowanych i kontrolowanych warunkach. Wśród metod instrumentalnych są też takie, które polegają na obserwacji i rejestracji innych fizjologicznych korelatów emocji niż te, które są brane pod uwagę w klasycznym badaniu poligraficznym. Do takich metod należy np. obserwacja zmian temperatury twarzy, zmian w głosie, ruchu gałek ocznych. Ale do metod instrumentalnych zaliczamy także takie, które dokonują obserwacji i oceny zmian nie na poziomie fizjologicznym, jak badanie poligraficzne czy inne wspomniane metody, a na poziomie neurofizjologicznym. Do takich metod należą bezpośrednie obserwacje i rejestracje pracy mózgu za pomocą badania EEG lub fMRI.

Jak się okazuje, w literaturze polskiej nawet nazwa „poligraf” bywa kwestionowana, próbuje się ją zastąpić nazwą „wariograf”.

Te niezbędne dla spójności całego projektu ustalenia terminologiczne zostały dokonane w trzech pracach:

- J. Widacki, K. Dukała, *Detekcja kłamstwa – czyli czego?*, „Problemy Kryminalistyki” 2015, nr 287 (1), s. 3–16;
- J. Widacki, *Poligraf czy wariograf?*, „Problemy Kryminalistyki” 2016, nr 294 (4), s. 15–19;
- J. Widacki, *The first polygraph?*, „European Polygraph” 2016, 3 (37), s. 111–116.

ANEKS A

Jan Widacki, Karolina Dukała

Detekcja kłamstwa – czyli czego?*

Wprowadzenie

Klasyk kryminalistyki Hans Gross twierdził, że „większość pracy kryminalistyka to walka z kłamstwem”¹. Nic więc dziwnego, że w literaturze kryminalistycznej i psychologiczno-sądowej często używane są nazwy, „detekcja kłamstwa”, „instrumentalna detekcja kłamstwa”, „nieinstrumentalna detekcja kłamstwa” (lie-detection)². W literaturze spotyka się również takie określenia jak „wykrywanie oszustw”³ a ściślej: „oszukiwania” (detection of deception)) czy „psychologiczna ocena wiarygodności zeznania”⁴. Niezależnie od tego, w języku opinii psychologiczno-sądowych używana jest nazwa „wiarygodność psychologiczna”, będąca jakimś kryterium oddzielania wypowiedzi szczerych i kłamliwych.

Już w XIX wieku Pierre Simone de Laplace zaproponował podział na różne kategorie zeznań. Wy różnił zeznania prawdziwe czyli takie, które są dokładnym odzwierciedleniem rzeczywistości oraz zeznania szczerze – które są subiektywnie prawdziwe dla świadka (tj. świadek ma intencję zeznawać prawdziwie, zgodnie z tym co pamięta), ale mogą nie odzwierciedlać obiektywnie istniejącej rzeczywistości. Te kryteria pozwalają na utworzenie czterech typów kategorii zeznań:

- prawdziwych i szczerych (odzwierciedlających obiektywną rzeczywistość oraz subiektywnie prawdziwych);
- nieprawdziwych i nieszczerych (świadek dąży do tego, by przekazać fałszywe dane; które są niezgodne z obiektywną prawdą);
- nieprawdziwych, ale szczerych (świadek stara się zeznawać zgodnie ze swoją wiedzą i jest przekonany, że mówi prawdę, ale jego wypowiedź np. na skutek błędów pamięciowych, zapomnienia itd. jest niezgodna z prawdą obiektywną) oraz
- nieszczerych, ale prawdziwych (bardzo rzadko występująca grupa zeznań: świadek chce skłamać i zeznaje niezgodnie z tym, co pamięta, ale mimo to przekazuje informacje zgodne z obiektywną prawdą).

* J. Widacki, K. Dukała, *Detekcja kłamstwa – czyli czego?*, „Problemy Kryminalistyki” 2015, 287 (1), s. 3–16; wersja anglojęzyczna *Lie detection - of what?*, „Problemy Kryminalistyki” 2015, 287 (1), s. 61–74.

¹ H. Gross, *Criminal psychology: A manual for judges, practitioners and students*, Little, Brown and Comp., Boston 1918, s. 474.

² Por. np.: C.D. Lee, *The instrumental detection of deception*, Ch. Thomas, Springfield 1953; A. Vrij, *Detecting Lies and Deceit. The psychology of lying and implications for professional practice*, John Wiley & Sons, Chichester, 2000; J. Ulatowska, *Knowledge of cues to deception – looking for its determinants*, *Problems of Forensic Sciences* 2009, 80, s. 411–428; *Kryminalistyka*, J. Widacki (red.), C.H. Beck, Warszawa 2012, s. 106–109, s. 393–395.

³ K. Cantarero, *Wykrywanie kłamstwa w komunikacji interpersonalnej*, *Psychologia Społeczna* 2009, 4, s. 167–176.

⁴ B. Wojciechowski, *Psychologiczna analiza treści zeznań świadków i ocena ich wiarygodności*, *Palestra* 2012, nr 1–2, s. 70–80.

Większość obecnie stosowanych metod „detekcji kłamstwa” odnosi się zatem do kategorii wypowiedzi subiektywnie nieszczerych i obiektywnie nieprawdziwych. Spotkać można też twierdzenia, że przedmiotem metod detekcji nie jest kłamstwo czy szczerłość, ale „świadomość winnego” (*guilty knowledge*)⁵ czy wprost „wina” (*guilty*)⁶, „wiedza o realiach czynu”⁷ albo „związek emocjonalny badanego z czynem” lub „ukrywanie informacji” (*concealed information*)⁸. Zauważyć jednak wypada, że ilekroć mowa jest o „świadomości winnego”, czy „wiedzy o realiach czynu” albo „związku emocjonalnym badanego z czynem”, zawsze będzie to ukrywana (zatajana) świadomość winnego, zatajana wiedza o czynie, zatajany związek emocjonalny z czynem.

Zatem na pierwszy rzut oka widać, że termin „kłamstwo” jako przedmiot detekcji bądź oceny używany jest w literaturze psychologicznej i kryminalistycznej w różnorodnych znaczeniach, w dodatku znaczeniach nie do końca określonych. Uporządkowanie terminologii wydaje się więc konieczne.

1. Kwestie językowe

Wedle współczesnego słownika języka polskiego, „kłamstwo” to twierdzenie niezgodne z rzeczywistością, wypowiedziane z zamiarem wprowadzenia kogoś w błąd⁹. Ale wedle tego słownika, „kłamstwo” to także „fałsz”, „nieprawda”. W słowniku języka polskiego z 1961 roku¹⁰ dodane jest jeszcze jedno znaczenie słowa „kłamstwo”, a ściślej jego kolejny synonim: „kłamanie”. Słownik ten wprowadza też istotne, jak się okaże dla dalszych rozważań, rozróżnienie na „kłamstwo negatywne” i „kłamstwo pozytywne”¹¹. Z tym pierwszym, nazwanym „negatywnym” mamy do czynienia, na przykład, gdy ktoś zapytany przez kogoś nie przyznaje się do tego co zrobił. Z „kłamstwem pozytywnym” mamy do czynienia wówczas, gdy ktoś zmyśla jakąś rzecz, która faktycznie nie miała miejsca. Z kolei współczesne słowniki synonimów języka polskiego dla słowa „kłamstwo” wymieniają sto kilkadziesiąt synonimów, grupując je w ponad 40 najważniejszych grupach znaczeniowych.

Zdaje się nie ulegać wątpliwości, że istotą kłamstwa jest wypowiedź świadomie przekazująca informacje nieprawdziwe. Przekazanie informacji nieprawdziwych może przyjąć dowolną formę komunikacji. Może mieć postać zdania wypowiedzianego ustnie, bądź zdania w formie pisemnej. Może też przyjąć formę znaku (w tym mimicznego, pantomimicznego) lub gestu. W przypadku gdy wypowiedź przekazująca informacje nieprawdziwe, ma formę zdania (mówionego lub pisanego) jest zdaniem logicznie fałszywym. Przy czym wypowiadający je (lub piszący) ma tego świadomość, lub co najmniej, nie jest pewien jaka jest logiczna wartość wypowiadanego (pisanego) zdania, ale on przekazuje je tak, jakby był pewien, że przekazuje zdanie prawdziwe.

Tak więc warunkiem koniecznym zaistnienia kłamstwa, jest świadome wypowiedzenie zdania fałszywego. Czy jest to zarazem warunek wystarczający? Czy każde wypowiedzenie zdania fałszywego jest kłamstwem? Czy może tylko takie, które jest do kogoś adresowane, czyli ma odbiorcę, co więcej, jest wypowiedziane w celu wprowadzenia tego adresata-odbiorcy w błąd?

Inaczej mówiąc, czy Kowalski kłamie, mówiąc w pustym pokoju, do lustra oczywiście nieprawdziwe zdanie: „*jestem Napoleonem Bonaparte*”? A co wtedy, gdy tenże Kowalski przy kolacji powie do żony, która zna go od kilkudziesięciu lat, a poza tym zna historię, „*jestem Napoleonem Bonaparte*”? W obydwu przypadkach Kowalski wypowie zdanie fałszywe. Czy uznamy to za kłamstwo, czy raczej za jakiś żart, wygłup? Czy zgodnie z intuicją języka polskiego uznamy Kowalskiego za kłamcę, czy raczej za żartownisia, ewentualnie błazna?

⁵ D.T. Lykken, *The validity of the guilty knowledge technique*, Journal of Applied Psychology 1960, 44, 258–262; D.T. Lykken, *Guilty knowledge test – the right way to use lie-detector*, Psychology Today 1975, 8, 10, s. 56–60.

⁶ D.T. Lykken, *The GSR In detection of guilt*, Journal of Applied Psychology 1959, 43, s. 385–388.

⁷ M. Kulicki, *Kryminalistyka. Wybrane problemy teorii i praktyki śledczo-sądowej*, Toruń 1994, s. 482.

⁸ Por. M. Nakayama, *Practical use of the Concealed Information Test for „criminal investigation In Japan*, [w:] M. Kleiner (red.), *Handbook of polygraph testing*, Academic Press, London 2002, s. 49–86.

⁹ *Podręczny słownik języka polskiego*, oprac. E. Sobol, Wyd. Naukowe PWN, Warszawa 1996, s. 334.

¹⁰ W. Doroszewski (red.), *Słownik języka polskiego*, PAN, Warszawa 1961, t. III, lit. H-K, s. 746.

¹¹ *Ibidem*.

Kolejny problem, jaki się tu wyłania, to problem, czy mylący świadomie gest, nie będący ex-definitione zdaniem (tym bardziej zdaniem w sensie logicznym) jest kłamstwem? Gdy ktoś pyta sekretarkę, czy dyrektor jest u siebie w gabinecie i on tam rzeczywiście jest, a sekretarka odpowiada przeczącym ruchem głowy, wprowadzając pytającego w błąd, to co: skłamała czy nie-skłamała?

Jeśli uznamy, że skłamała to wynikają z tego dwa istotne wnioski:

- 1) Kłamanie nie wymaga wypowiedzenia pełnego zdania fałszywego;
- 2) Istotą kłamania jest chęć wprowadzenia odbiorcy komunikatu w błąd, a nie forma w jakiej komunikat jest przekazany;

Gest (skiniecie głową, przeczący ruch głowy lub ręki, bezradne rozłożenie rąk itp.) zastępują w procesie komunikacji słowa „tak”, „nie”, „nie wiem”. Czy taka wypowiedź gestem, lub jednowyrazowa („tak”, „nie”) może być z logicznego punktu widzenia prawdziwa lub fałszywa? Sama taka wypowiedź oczywiście nie, ale w kontekście pytania, może mieć wartość logiczną.

Jeśli na pytanie „czy dyrektor jest u siebie w gabinecie” sekretarka odpowiada słowem „tak” lub „nie”, albo gestem zastępującym te słowa, to faktycznie przekazany komunikat ma postać zdania: „dyrektor jest u siebie w gabinecie” lub „dyrektora nie ma w gabinecie”, a te zdania mają wartość logiczną.

2. Kwestie filozoficzne

Z dotychczasowych rozważań wynika, że **kłamstwem jest przekazanie odbiorcy komunikatu który jest zdaniem fałszywym, w intencji, by odbiorca uznał go za zdanie prawdziwe.**

Dla filozofii kłamstwo jest przede wszystkim problemem etycznym. Czy kłamstwo (lie) jest specyficznym złem, które człowiek wprowadził do przyrody¹², czy też występuje ono w przyrodzie choćby w formie mimetyzmu (upodobniania się do otoczenia lub innych gatunków), czy mimikry (upodobnianie się organizmów bezbronnych, do organizmów zdolnych do obrony), częściej u zwierząt a nawet roślin, czy bardziej wyrafinowanych zachowań zwierząt mających wprowadzić w błąd przeciwnika czy ofiarę. Gdyby te zachowania uznać za formę kłamstwa, kłamstwo ludzkie wyróżniałoby się jedynie formą (językiem), bardziej czasem wyrafinowanym celem i stopniem świadomości.

Innym problemem etycznym jest zakres usprawiedliwionego kłamstwa. Ocena etyczna nie jest związana tyle z samym kłamstwem, co raczej z celem dla którego zostało użyte. O usprawiedliwionym etycznie kłamstwie można mówić w medycynie¹³ gdy na przykład nie chce się pozbawiać nadziei chorego i zataja się przed nim faktyczny stan jego zdrowia. Kłamstwo może być usprawiedliwione także w polityce, gdy na przykład dezinformuje się celowo wywiad obcego państwa, gdy agent wywiadu udaje, że jest dyplomatą itp. Kłamstwo może być usprawiedliwione w życiu codziennym, na przykład względami grzecznościowymi, gdy mówimy komuś niezasłużony, ale i oczekiwany komplement. Etycznie usprawiedliwione może być kłamstwo w wychowaniu dzieci, gdy na przykład małemu dziecku nie mówimy, że jego ojciec zmarł albo trafił do więzienia, tylko mówimy –oczywiście nieprawdziwe zdanie – że ojciec wyjechał. Na dobrą sprawę kłamiemy też, opowiadając dziecku bajki, mówiąc, że prezent przyniósł mu św. Mikołaj itp. Usprawiedliwione może być także kłamstwo w biznesie, gdy na przykład w negocjacjach posługujemy się bluffem. Usprawiedliwionemu kłamstwu poświęcona jest wydana niedawno monografia J. Kucharskiego¹⁴.

Wydaje się, że większym problemem, niż kłamstwo, uznane przez Duparta¹⁵ za „najstraszniejsze zło społeczne”, jest perspektywa całkowitego wyeliminowania kłamstwa, jaką stwarzają współczesne badania mózgu i ich jak się wydaje nie tak odległe w czasie konsekwencje, między innymi takie, że

¹² M. Buber, *Good and Evil*, Charles Scribner's Son, New York, s. 7.

¹³ A. Szczeklik, *Prawda dla ciężko chorego*, Diametros 2005, 4, s. 151; E. Kubler-Ross, *Rozmowy o śmierci i umieraniu*, Media Rodzina, Poznań 1998, 55–65; por. też J. Kucharski, *Usprawiedliwione kłamstwo we współczesnej etyce stosowanej*, Akademia Ignatianum – Wydawnictwo WAM, Kraków 2014, s. 227 i nast.

¹⁴ J. Kucharski, *Usprawiedliwione kłamstwo we współczesnej etyce stosowanej*, Akademia Ignatianum, Wydawnictwo WAM, Kraków 2014; por. też: J. Widacki, A. Szuba-Boroń: Book Review: J. Kucharski, *op. cit.*, *European Polygraph*, 2014, 4, s. 199–203.

¹⁵ G.L. Dupart, *Kłamstwo. Studium psycho-socjologiczne*, Gebethner i Wolf, Warszawa 1905, s. 9–10.

będzie możliwe czytanie ludzkich myśli, co nie tylko wyeliminuje kłamstwo, ale przyniesie koniec naszej cywilizacji w jej obecnym kształcie¹⁶. Pozbawi nas bowiem nie tylko wszelkiej prywatności, ale zmusi do ujawnienia prawdy, także tam, gdzie kłamstwo jest etycznie usprawiedliwione i potrzebne.

Dla filozofii kłamstwo jest jednak nie tylko problemem etycznym, ale także problemem epistemologicznym. Próbuując zdefiniować kłamstwo, chcąc nie chcąc, dochodzimy do problemu prawdy. Pojęcie prawdy jest przedmiotem rozważań zarówno logiki (zwłaszcza semantyki logicznej) jak i filozofii, w której jest centralnym problemem epistemologii.

Z filozoficznego punktu widzenia, prawda jest cechą naszego poznania, a nie właściwością rzeczywistości. Inaczej mówiąc prawda odnosi się nie do faktów, tylko do sądów o faktach, do zdań, które o faktach są wypowiedzane.

Filozofia zachodnia próbuje ustalić czym jest prawda, jak się ją stwierdza i wedle jakich kryteriów. Propozycji jest wiele. Jedna z nich, historycznie najstarszą, ale wciąż aktualną, jest klasyczna koncepcja prawdy. Wedle tej propozycji (zwanej czasem teorią), prawda jest to zgodność rzeczy i umysłu (*veritas est adequatio rei et intellectus*). Czyli zgodność rzeczywistości i sądu o tej rzeczywistości. Arystoteles, uważany za prekursora tej koncepcji w „Metafizyce” pisał, że „Fałszem jest powiedzieć o tym co jest, że nie jest, lub o tym co nie jest, że jest; prawdą jest powiedzieć o tym co jest, że jest, lub o tym co nie jest, że nie jest”¹⁷.

Na pierwszy rzut oka wszystko tu jest oczywiste i zgodne z intuicją. Problem zaczyna się, gdy chcemy określić, na czym owa „zgodność” (*adequatio*) rzeczy i sądu o rzeczy polega. Czy w szczególności realny stan rzeczy jest podobny do myśli o nim? Czy chodzi tu o dosłowną adekwatność, czy o przekonanie, że ta adekwatność istnieje? Trudności jakie rodzi klasyczna koncepcja prawdy próbują przełamać tzw. nieklasyczne koncepcje prawdy. Jedną z nich jest koncepcja oczywistości, zwana też koncepcją ewidencji. Twierdzi ona, że prawdziwy jest taki sąd o rzeczywistości, który jest oczywisty. Neokantysta Heinrich Rickert (1863–1936), twierdził, że twierdzenie oczywiste narzuca się z koniecznością, odczuwalną jako przymus, czy też obowiązek przyjęcia go¹⁸. Inną koncepcją nieklasyczną, jest koncepcja zgody powszechnej. Jeszcze inną koncepcją nieklasyczną, jest koncepcja koherencyjna. Wedle tej koncepcji, prawdziwość poprawnie zbudowanego zdania (sądu) polega na jego zgodności z pozostałymi zdaniami, uznanymi już za prawdziwe. Chodzi więc o zgodność logiczną, polegającą na niesprzeczności w ramach jakiegoś systemu.

Jeszcze inna koncepcja prawdy, jest koncepcja utylitarystyczna (pragmatyczna). Wedle tej koncepcji, prawdziwość polega na tym, że jej przyjęcie prowadzi do pożytecznych, praktycznych celów. Jeśli uznanie za prawdziwe zdania (sądu) lub teorii jest pożyteczne, praktyczne, to zdanie to (teoria) jest prawdziwe.

3. Kłamstwo w ujęciu psychologicznym

W psychologii kłamstwo jest definiowane na różne sposoby. W psychologii rozwoju można wyróżnić trzy podejścia dotyczące kłamstwa: **zdaniowe** (kłamstwo jest wtedy, gdy jedna osoba przekazuje fałszywą informację drugiej, w celu wprowadzenia jej w błąd¹⁹; **prototypowe** (każdą informację można umieścić na kontinuum, którego jeden kraniec stanowi prototypowe kłamstwo, a drugi prototypowa prawda: prawdziwy komunikat, który nadawca przekazuje drugiej osobie z intencją poinformowania jej²⁰ oraz tzw. **model folklorystyczny kłamstwa**²¹ w którym zakłada się, że kłamstwem jest dopiero

¹⁶ J. Widacki, *Pożegnanie współczesności – nauka i technika przeciw terroryzmowi. Terror nauki i techniki*, [w:] G. Babiński, M. Kapiszewska (red.), *Zrozumieć współczesność. Księga jubileuszowa prof. H. Kubiaka*, Oficyna Wydawnicza AFM, Kraków 2009, s. 287–294.

¹⁷ A. Węgrzecki, *Zarys filozofii*, Wydawnictwo AE, Kraków 2005, s. 108.

¹⁸ A. Węgrzecki, *op. cit.*, s. 110.

¹⁹ R.M. Chisholm, T.D. Feehan, *The intent to deceive*, *Journal of Philosophy* 1977, 74, s. 143–159.

²⁰ L. Coleman, P. Kay, *Prototype semantics: The English word lie*, *Language* 1981, 57, s. 26–44.

²¹ E.E. Sweetser, *The definition of lie: An examination of the folk models underlying a semantic prototype*, [w:] D. Holland (red.), *Cultural models in language and thought*, s. 43–66, Cambridge University Press, New York 1987.

taki komunikat, którego celem jest skrzywdzenie drugiej osoby. Tym samym kłamstwo altruistyczne (na przykład niezasłużone komplementowanie drugiej osoby, w celu zrobienia jej przyjemności) nie są kłamstwami wedle tego założenia.

W psychologii komunikacji zakłada się, że kłamstwo, to rodzaj komunikatu nadany przez nadawcę do odbiorcy oraz wynik oddziaływania na siebie partnerów tej interakcji²². Innymi słowy, w kłamstwie można wyróżnić wszystkie aspekty procesu komunikacji można wyróżnić wszystkie aspekty procesu komunikacji, a więc nadawcę, odbiorcę, kod, kontekst, przesyłaną informację (kłamstwo samo w sobie) oraz informację zwrotną. Jak wskazuje Cantarero²³ badania nad kłamstwem mogą dotyczyć wszystkich tych aspektów. I tak, badania dotyczące kontekstu dotyczą wpływu otoczenia, w tym innych osób w pomieszczeniu na częstotliwość czy jakość kłamania, badania nad nadawcą i odbiorcą skupiają się nad wyznaczeniem cech osobowości, temperamentu, czy stosowanych strategii kłamcy i osoby okłamywanej, które wpływają na skuteczność kłamania. Badania nad kodem (np. słowa, gesty, ruch) i komunikatem z kolei dotyczą metod, dzięki którym kłamstwo można wykryć.

Psychologowie zajmujący się kłamstwem w kontekście społecznym lub sądowym zwykle określają je jako fałszywy komunikat, którego celem jest przyniesienie korzyści osobie oszukującej²⁴, lub bardziej precyzyjnie, czyn, którego celem jest wytworzenie w innej osobie przekonania co do którego kłamiący jest przekonany, że jest fałszywe²⁵. Tym samym, na przykład fałszywe wspomnienia (na przykład, że było się molestowanym, podczas gdy nie miało to miejsca) nie są traktowane jako oszustwo (kłamstwo).

4. Kłamstwo jako przedmiot badania w kryminalistyce i psychologii sądowej

Przedmiotem zainteresowania kryminalistyki i psychologii sądowej jest zarówno kłamstwo pozytywne, jak i kłamstwo negatywne. Kłamstwo negatywne polega na negowaniu rzeczywistości, na przykład wtedy, gdy podejrzany nie przyznaje się do popełnienia swego czynu (czynu, który popełnił). Z kłamstwem pozytywnym mamy do czynienia, gdy np. ktoś kogoś fałszywie oskarża. Można sobie bez trudu wyobrazić sytuacje, gdy mamy do czynienia równocześnie z kłamstwem negatywnym i pozytywnym równocześnie: np. ktoś nie przyznając się do czynu, którego dokonał, obciąża nim kogoś innego. **W praktyce interesują nas tylko sytuacje, gdy ktoś świadomie wypowiada zdania fałszywe w intencji, by ktoś inny uznał je za prawdziwe.** Może tu chodzić zarówno o wypowiedzi, które mają postać zdania w sensie logicznym, jak i takie, które dopiero z kontekstem pytania tworzą zdanie w sensie logicznym i to zdanie fałszywe, a są pojedynczymi wyrazami (tak-nie) lub zastępującymi je gestami. Rodzi się pytanie, czy dysponujemy dziś technikami (instrumentalnymi lub nieinstrumentalnymi) pozwalającymi na wykrycie tak rozumianego kłamstwa? Trudno udzielić jednoznacznej odpowiedzi na to pytanie. **Jak dotąd jest jedynie sens podejmować próby wykrycia kłamstwa (spełniającego podane wyżej warunki), z którego wykryciem łączą się jakieś konsekwencje dla kłamiącego.** Dzieje się tak, ponieważ dotychczasowe metody detekcji kłamstwa pozwalają jedynie na wykrycie fizjologicznych i poznawczych korelatów emocji, które mają towarzyszyć kłamstwu. Z ich istnienia, wnioskujemy o kłamstwie lub jego braku. Temu zagadnieniu trzeba poświęcić nieco więcej uwagi.

Krótko mówiąc, nie wykrywamy kłamstwa jako takiego, w każdym razie nie wykrywamy go bezpośrednio. Wykrywamy towarzyszące mu zjawiska (np. fizjologiczne korelaty emocji), które obserwujemy bezpośrednio. Z ich istnienia, lub nieistnienia wnioskujemy o emocjach, a z istnienia tych ostatnich, w kontekście badania (treści wywiadu przed testowego, treści pytań testowych) wnioskujemy o kłamstwie lub jego braku. Tak wygląda detekcja kłamstwa w przypadku badań poligraficznych, lub technik nieinstrumentalnych, polegających na obserwacji zmian behawioralnych

²² C.H. White, J.K. Burgoon, *Adaptation and communicative design: Patterns of interaction in truthful and deceptive conversations*, Human Communication Research 2001, 27, s. 9–37.

²³ K. Cantarero, *Wykrywanie...*, op. cit.

²⁴ R.W. Mitchell, *A framework for discussing deception*, [w:] R.W. Mitchell, N. S. Mogdil (red.), *Deception: Perspectives on human and nonhuman deceit*, s. 3–40, State University of New York Press, Albany 1986.

²⁵ A. Vrij, *Detecting...*, op. cit.

(głównie ruchów wyrazowych: mimicznych, pantomimicznych). Co innego jest natomiast przedmiotem detekcji w przypadku technik bazujących na analizie treści wypowiedzi.

5. Techniki badań poligraficznych i przedmiot ich detekcji

Badania poligraficzne są rutynowo stosowaną od lat 20. XX wieku instrumentalną techniką detekcji kłamstwa. Polegają one na wywołaniu poprzez właściwie przeprowadzony wywiad przedtestowy, odpowiednio sformułowane pytania o określonej strukturze, ułożone dodatkowo w kolejności wedle pewnego schematu, a następnie obserwowanie i rejestrowanie za pomocą aparatu (poligrafu) fizjologicznych korelatów emocji wywoływanych przez pytania testu.

5.1. Technika klasyczna

Technika klasyczna, zwana też techniką Keelera, albo technika pytań „związane-niezwiązane” jest najstarszą z rutynowych technik badań poligraficznych. Została ona opracowana ostatecznie w latach 30. XX wieku przez Leonarda Keelera²⁶. Technika ta posługuje się testami znającymi dwa rodzaje pytań: związanych ze sprawą (relevant questions) i niezwiązanych ze sprawą (irrelevant questions). Te pierwsze pytania nazywa się czasem pytaniami krytycznymi, te drugie – obojętnymi.

Pytania, jak wszystkie w badaniach poligraficznych muszą być tak sformułowane, by badany odpowiadać mógł na nie krótko: „tak” lub „nie”. A zatem z logicznego punktu widzenia muszą to być „pytania rozstrzygnięcia”²⁷.

Pytania związane (krytyczne, relevant questions) dotyczyć mają zdarzenia będącego przedmiotem postępowania i ewentualnej roli badanego w tym zdarzeniu. Tak np. w sprawie o zabójstwo kobiety, takimi typowymi pytaniami są pytania takie jak: „czy to ty zabiłeś kobietę?”, „czy wiesz, kto zabił tę kobietę?”, „czy w dniu zabójstwa byłeś na tym miejscu?”, i t.p. Zakłada się, że badany, który nie przyznaje się do udziału w zabójstwie, na pytania te odpowiadał będzie „nie”.

Pytania niezwiązane (obojętne, irrelevant questions) dotyczyć muszą sytuacji niezwiązanych ze zdarzeniem będącym przedmiotem dociekań. Takimi pytaniami, w sprawie o zabójstwo kobiety, będą pytania „czy dziś jest wtorek?” (o ile rzeczywiście jest wtorek), „czy znajdujemy się w Krakowie?” (o ile badanie rzeczywiście realizowane jest w Krakowie), „czy masz na imię Józek” o ile badany rzeczywiście ma na imię Józek i nie próbuje ukryć swej tożsamości. Pytania te są tak sformułowane, by nie były same przez się emocjogenne, i by badany mógł na nie odpowiedzieć „tak”.

Technika ta zakłada, że badany odpowiadając „tak” na pytania niezwiązane, będzie mówił prawdę, zaś odpowiadając „nie” na pytania związane (krytyczne), o ile był sprawcą lub miał inny udział w zdarzeniu będącym przedmiotem badania, kłamał. Ocena, czy odpowiadając „nie” na pytanie związane (krytyczne) badany kłamał, jest zadaniem badającego.

Technika klasyczna dopuszczała obok testów klasycznych, stosować dodatkowo tzw. „testy szczytowego napięcia” (peak of tension, POT). Istota takiego testu polega na tym, że pytanie krytyczne ukryte jest wśród innych pozornie podobnych pytań²⁸. Na przykład w sprawie, gdy zabójstwa dokonano przy użyciu noża, a badany twierdzi, że nie tylko nie dokonał tego czynu, ale w dodatku nie wie, ani się nie domyśla jakim narzędziem zabójstwa dokonano, można mu zadać pytania następujące:

1. Czy uderzyłeś tę kobietę łomem? (obojętne)
2. Czy zarąbałeś tę kobietę siekierą? (obojętne)
3. Czy uderzyłeś tę kobietę wazonem? (obojętne)
4. Czy ugodziłeś tę kobietę nożem? (krytyczne)

²⁶ J. Widacki, *Wprowadzenie do problematyki badań poligraficznych*, Wyd. MSW, Warszawa 1981, s. 56.

²⁷ Por. *Mała encyklopedia logiki*, Ossolineum, Wrocław–Warszawa–Kraków 1970, s. 233.

²⁸ J. Widacki, *Wprowadzenie do problematyki badań poligraficznych*, Warszawa 1981, s. 62; J.E. Reid, F.E. Inbau, *Truth and Deception. The polygraph (lie-detector) technique*, Williams & Wilkins Comp., Baltimore 1977, s. 55.

5. Czy strzeliłeś do tej kobiety? (obojętne)
6. Czy udusiłeś tę kobietę sznurem?(obojętne)

Zakłada się, że sprawca czynu, i w zasadzie tylko on, będzie wiedział, które pytanie jest krytyczne, i odpowiadając na to jedno pytanie słowem „nie” – skłamię. Zadaniem badającego jest zatem, podobnie jak przy testach klasycznych ustalenie, czy odpowiadając „nie” na pytanie krytyczne, badany kłamał.

5.2. Techniki pytań kontrolnych

Twórcą techniki pytań kontrolnych, opartych o testy pytań kontrolnych, jest John Reid²⁹. Reid do badań poligraficznych wprowadził, obok pytań związanych (krytycznych) i niezwiązanych (obojętnych), trzeci rodzaj pytań: pytania kontrolne (control questions). Są to pytania tematycznie dotyczące rodzaju przestępstwa będącego przedmiotem badania, ale tak sformułowane, by badany odpowiedział na nie „nie”, ale odpowiadając albo świadomie skłamał, albo przynajmniej nie był pewien, czy nie skłamał. Klasyczny układ pytań w teście pytań kontrolnych przez Reida jest następujący³⁰:

1. Czy nazywają cię Rudy?[w czasie wywiadu przedtestowego badany powiedział, że zwykle nazywają go „Rudy”]
2. Czy masz ukończone 21 lat? [gdy rzeczywiście badany ma ukończone 21 lat, nie ukrywa tego, czy zataja]
3. Czy ostatniej soboty strzelałeś do Johna Jonsa?
4. Czy znajdujemy się teraz w Chicago? [o ile badanie odbywa się w Chicago];
5. Czy to ty zabiłeś Johna Jonsa?
6. Czy oprócz tego co powiedziałeś, czy ty ukradłeś jeszcze coś?
7. Czy ty chodziłeś do szkoły?
8. Czy ty ukradłeś zegarek Johna Jonsa w nocy w ostatnią sobotę?
9. Czy wiesz, kto strzelał ostatniej soboty do Johna Jonsa?
10. Czy ty kiedykolwiek ukradłeś coś w miejscu pracy?

W teście tym, pytania 1,2, 4, 7 to pytania obojętne. Pytania 3,5, 8,9 to pytania krytyczne, pytania 6 i 10, to pytania kontrolne.

W technice badań opartych na testach pytań kontrolnych, nie porównuje się reakcji na pytania krytyczne (związane) z reakcjami na pytania obojętne (niezwiązane), jak to ma miejsce w przypadku techniki opartej na testach klasycznych, ale porównuje się reakcje na pytania krytyczne z reakcjami na pytania kontrolne. Reakcję na pytania 3 i 5 porównuje się z reakcjami na pytanie 6, reakcje na pytania 8 i 9 z reakcjami na pytanie 10.

Technika Reida zakładała wykonanie dwóch identycznych testów pytań kontrolnych ułożonych wedle powyższego schematu, nadto testu „pytań mieszanych”, zawierających wszystkie te same pytania, co testy pytań kontrolnych, ale zadanych w innej nieznanej uprzednio badanemu kolejności. W zależności od potrzeb i możliwości, w ramach tej techniki badań wykonywano dodatkowo testy typu POT a także testy pomocnicze „na tak” (Yes-test) lub testy pytań kontrolnych w „wersji milczącej” (Silent Answer Test, SAT). W pierwszym z tych testów, prosiło się badanego, by bez względu na to, czy będzie to odpowiedź prawdziwa czy fałszywa, na każde pytanie odpowiadał „tak”. Wedle zalecenia Reida-Inbaua³¹ taki test przeprowadza się wówczas, gdy testy podstawowe pytań kontrolnych, łącznie z testem pytań mieszanych nie dały podstaw do rozstrzygnięcia, a istnieje podejrzenie, że badany w czasie trwania badania próbował manipulować swymi reakcjami. Zakłada się, że taka osoba, w te-

²⁹ J.E. Reid, *A revised questioning technique In lie-detection test*, Journal of Criminal Law, Criminology and Police Science, 1947, 37, s. 542–547.

³⁰ J.E. Reid, F.E. Inbau, *Truth...*, op cit., s. 31.

³¹ *Ibidem*, s. 32–33.

ście „na tak” będzie próbowała sztucznie wywołać u siebie wyraźną reakcję na pytania krytyczne by w ten sposób zmylić badającego. W teście „w wersji milczącej” badany jest instruowany, by na pytania testowe nie odpowiadał głośno, tylko aby odpowiadał w myślach, zgodnie z prawdą. Test „milczących odpowiedzi” jak wynikało z praktyki, dawał, przy niektórych typach osobowości lepsze wyniki niż zwykły test pytań kontrolnych, przy którym badany musi głośno mówić „tak” lub „nie”.

Przy technice pytań kontrolnych Reida, jako pomocnicze kryteria diagnostyczne, obok oceny reakcji zapisanych przez poligraf, po uwagę brano pomocniczo ocenę ewentualnych prób zakłóceń zapisów, werbalne i niewerbalne symptomy towarzyszące zachowaniu badanego przed, w czasie i po badaniu³².

W oparciu o wszystkie przesłanki: reakcje na pytania krytyczne (w relacji do reakcji na pytania kontrolne), reakcje w testach POT, w testach pomocniczych, ocenę ewent. zaburzeń zapisu, ocenę zachowania (poprzez pryzmat werbalnych i niewerbalnych symptomów), badanego zaliczano do jednej z trzech grup: nieszczerych („deceptive”), szczerych (non-deceptive) lub grupy o braku rozstrzygnięcia (inconclusive).

W prawidłowej metodologicznie opinii z badań należało to zapisać: *„badany reaguje na pytania krytyczne testów, tak, jak zwykle reagują osoby, które na pytania takie odpowiadają nieszczerze, to jest kłamliw, lub zatajają fakt posiadania informacji związanych ze sprawą”*.

O ile odpowiedzi na poszczególne pytania krytyczne testów pytań kontrolnych (wraz z kontekstem pytania) mogą być zdaniem fałszywymi z logicznego punktu widzenia, a zatem ich wypowiedzanie do badającego, w intencji, aby zdania te uznać za prawdziwe, są kłamstwami, to **w badaniu nie ocenia się reakcji na każde z pytań z osobna, ale ocena szczerości badanego dokonana zostaje całościowo, na podstawie reakcji na wszystkie pytania. Nie wykrywa pojedynczych kłamstw na poszczególne pytania krytyczne, ale generalizuje ocenę na całość testu, kwalifikując badanego jako „nieszczerego” lub „szczerego” (albo, przy braku jednoznacznych reakcji, uznając wynik badania za nierozstrzygnięty)**. Nie diagnozuje się np.: „badany skłamał na pytanie 5”, „badany nie skłamał na pytanie 9”, ale wszystkie reakcje z testu uśrednia się zaliczając badanego do jednej w wskazanych wyżej grup: „szczerzy” (non-deceptive NDI), „nieszczerzy” (deceptive DI) lub „brak rozstrzygnięcia” (inconclusive INC). W dodatku, w „testach milczących” które pomocniczo też są wykorzystywane przy diagnozie, w ogóle nie mamy do czynienia z kłamstwem w wyżej opisanym znaczeniu, badany bowiem nic nie mówi, nic świadomie i celowo nie przekazuje.

Identycznie pod tym względem jest w technikach pytań kontrolnych powstałych z rozwinięcia techniki Reida, a więc z technice porównań strefowych Backstera, czy nowszych, takich jak : Federal Zone Comparison Test, Utah Zone Comparison Test³³.

Zatem w badaniach poligraficznych wykonywanych techniką pytań kontrolnych nie wykrywa się kłamstwa na poszczególne pytania testu, ale po sumie reakcji na poszczególne pytania krytyczne, porównanych z reakcjami na pytania kontrolne, wnosimy całościowo, czy badany w czasie badania jest szczerzy czy nieszczerzy, ściślej : czy stara się wprowadzać w błąd badającego czy nie³⁴. Dotyczy to zarówno metod jakościowych analizy zapisów, jak i metod ilościowych (numerycznych)³⁵.

Warto zwrócić uwagę, że w naukowej, fachowej literaturze amerykańskiej, co najmniej od lat 50. XX wieku, mówi się nie o kłamstwie (lie), jako przedmiocie badań poligraficznych, ale o „wprowadzaniu w błąd”, „nieszczerości”, „oszukiwaniu” (deception)³⁶. Przez „deception”, w języku angielskim

³² J. Widacki, *Analiza przesłanek diagnozowania w badaniach poligraficznych*, Wyd. UŚ, Katowice 1982, s. 78–84.

³³ J. Widacki (red.), *Kryminalistyka*, 2.wydanie, C.H. Beck, Warszawa 2012, s. 383.

³⁴ Por. J. Widacki (red.), *Badania poligraficzne w Polsce*, Oficyna Wydawnicza AFM, Kraków 2014, s. 196.

³⁵ Por. np.: J.A. Matte, *Forensic psychophysiology. Using the polygraph*, J.A.M. Publications, Williamsville, New York 1996, s. 398 i nast.; D.C. Raskin, Ch. Honts, J.C. Kircher (red.), *Credibility assessment. Scientific research and applications*, Elsevier – Academic Press, 2012, s. 4–7, 161–210, 305–307.

³⁶ Por. np.: C.D. Lee, *The instrumental detection of deception*, Ch. Thomas, Springfield, 1953; J.E. Reid, F.E. Inbau, *Truth...*, *op. cit.*; Ch.R. Honts, *Psychophysiological detection of deception*, *Current directions in psychological sciences*, 1994, s. 77–82; J. Vendemia, *Detection of deception*, *Polygraph* 2003, 32, s. 97–106; D.C. Raskin, Ch. R. Honts, J.C. Kircher (red.), *Credibility...*, *op. cit.*, s. 86–88.

rozumie się akt rozmyślnego spowodowania, aby ktoś uwierzył w coś, co nie jest prawdą³⁷, lub użycie oszustwa, podstęp, wprowadzenie w błąd³⁸ albo akt rozmyślnego dostarczenia lub pominięcia informacji w intencji wprowadzenia w błąd³⁹.

5.3. Technika Lykkena

W latach 1958 i 1959 Lykken przeprowadził dwa eksperymenty, których celem, jak podkreślał, było nie wykrycie „kłamstwa” ale „świadomości winnego” (guilty knowledge)⁴⁰.

W eksperymentach Lykkena, badanemu nie zadawano pytań takich jak przy badaniach technikami pytań kontrolnych, czy technika klasyczną. W każdym teście, a przeprowadzono ich kilka, zadawano jedno poprzedzone swoistym wstępem pytanie wprowadzające, czytając później alternatywne odpowiedzi, z których jedna była rzeczywiście związana z wyreżyserowanym zdarzeniem przestępnym. Pierwsze z tych pytań brzmiało:

„Jeżeli ty jesteś mordercą, to ty będziesz wiedział, co było niezwykłym przedmiotem prezentowanym w pokoju, w którym dokonano morderstwa. Czy był to:

1. magnetofon?
2. sztaluga?
3. bombonierka?
4. szachy?

Badany na te pytania nie odpowiadał, tylko ich słuchał, a psychogalwanometr rejestrował reakcje. Dla wzmocnienia motywacji badanych, dodatkowo założono im elektrody szokera elektrycznego, informując, że gdy zmiana GSR będzie wskazywała winę, dostana bolesne uderzenie prądem⁴¹. Jak widać w eksperymentach Lykkena badany w ogóle nie odpowiadał na pytania, zatem trudno by tu było mówić o „kłamstwie”(lie) czy „nieszczerości” (deception).

5.4. Przedmiot detekcji w badaniach poligraficznych wykonanych techniką pytań kontrolnych i techniką Lykkena

Zwykło się mówić, że techniki pytań kontrolnych zmierzają do wykrycia kłamstwa, techniki Lykkena (Guilty Knowledge Test, GKT; Concealed Information Test, CIT), do wykrycia „świadomości winnego”, „ukrytych informacji” lub „wiedzy o czynie”. Czy rzeczywiście, techniki te służą do wykrywania czegoś innego? Czy inny jest przedmiot rozpoznania?

W literaturze zwracano już uwagę, że zarówno w technikach pytań kontrolnych, jak i w technice **Guilty Knowledge Test przedmiotem detekcji jest w gruncie rzeczy to samo⁴². Jest nim różnie nazywana, nieszczerość badanego, diagnozowana po całym badaniu, a nie nieszczerość odpowiedzi na poszczególne pytania testu.**

Jaka jest różnica, gdy się powie: „badany odpowiadając na pytania krytyczne jest nieszczerzy” (jak mówimy w technikach pytań kontrolnych Control Question Test – CQT) czy gdy powiemy: „badany ukrywa swą wiedzę o faktach związanych ze sprawą” (jak mówimy w technikach wiedzy o czynie – Guilty Knowledge Test – GKT)?

³⁷ Oxford advanced learner's dictionary, Oxford Univ. Press, 2010, s. 391.

³⁸ The American Heritage dictionary, Houghton Mifflin Comp., Boston 1985, s. 371.

³⁹ D. Krapohl, S. Sturm, *Terminology reference for the science of psychophysiological detection of Deception*, APA, Chattanooga, Tenn., 1997, s. 20.

⁴⁰ D.T. Lykken, *The GSR...*, op cit.; D.T. Lykken, *The validity...*, op cit.

⁴¹ D.T. Lykken, *The GSR...*, op cit. Na tych eksperymentach opiera się późniejsza technika badania poligraficznego zwana techniką „Guilty Knowledge Test” (GKT) lub „Concealed Information Test (CIT).

⁴² J. Widacki, *W sprawie wyboru techniki badania poligraficznego. Czy technika oparta na testach GKT (CIT) jest lepsza od techniki opartej na testach CQT?*, Problemy Kryminalistyki 2011, 273, s. 5–10; J. Widacki, *Logical identity of conclusions from polygraph testing performed in Control Question Test (CQT) and Guilty Knowledge Test (GKT) techniques*, European Polygraph, 2011, 1 (15), s. 1–10.

W badaniach poligraficznych realizowanych dla potrzeb procesu karnego, ocenie z zasady podlega szczerść/nieszczerość odpowiedzi badanego dotyczących obszaru tematycznego określonego i ograniczonego pytaniami testu a nie szczerść/nieszczerość odpowiedzi na każde poszczególne pytanie. Inaczej jest w badaniach przed zatrudnieniowych (pre-employment), gdzie istotna może być ocena szarości odpowiedzi na poszczególne pytania.

6. Inne instrumentalne metody detekcji kłamstwa

Obok badań poligraficznych, w których stosowany jest aparat zwany poligrafem, rejestrujący i pozwalający na ocenę takich fizjologicznych korelatów emocji jak zmiany w przebiegu czynności oddychania, zmiany w przebiegu pracy serca, zmiany odruchu skórno-galwanicznego, ewentualnie także zmiany reakcji pletysmograficznych podejmowane były i są próby wykorzystania do detekcji kłamstwa innych aparatów, rejestrujących inne fizjologiczne korelaty emocji lub pozwalających na obserwację procesów zachodzących w mózgu.

Podejmowano próby detekcji kłamstwa w oparciu o obserwację zmian w brzmieniu głosu, stosując do tego specjalne urządzenia jak na przykład „Psychological Stress Evaluator” (PSE), „Voice Stress Analyzer” (VSA), czy „Computerized Voice Stress Analyzer” (CVSA)⁴³. Podejmowane są też próby detekcji kłamstwa w oparciu o ocenę zmian temperatury twarzy, rejestrowanych przez kamerę termowizyjną⁴⁴, albo ocenę ruchu gałek ocznych rejestrowane przez specjalnie skonstruowane do tego celu urządzenie. Wszystkie te metody zmierzały bądź zmierzają do celów identycznych z celami badań poligraficznych. Ich celem jest detekcja identycznie jak w badaniu poligraficznym rozumianego kłamstwa, o którego istnieniu wnioskuje się z oceny zmian fizjologicznych korelatów emocji, stymulowanych odpowiednio sformułowanymi pytaniami.

Inaczej wygląda instrumentalna detekcja kłamstwa oparta nie na wywoływaniu i obserwacji fizjologicznych korelatów emocji, ale na obserwacji procesów zachodzących w mózgu. Pierwszą z nich jest elektroencefalograficzna. Drugą, obserwacja mózgu metodą rezonansu magnetycznego (fMRI). Pierwsze eksperymentalne próby wykrywania kłamstwa przy pomocy EEG podejmowane były przez Obermanna jeszcze w końcówce lat 30⁴⁵, a później prowadzone w latach 70. XX wieku m.in. przez Dufka i Richtera⁴⁶, Guljajewa i Bychowskiego⁴⁷. W ostatnich latach, począwszy od końca lat 90. XX wieku powróciło zainteresowanie dla wykorzystanie badań EEG do detekcji kłamstwa⁴⁸.

W ostatnich kilkunastu latach wykonano wiele badań eksperymentalnych polegających na wykorzystaniu rezonansu magnetycznego (fMRI) do detekcji kłamstwa⁴⁹. Nie wdając się w tym miejscu w rozważania, na ile te badania mogą (i ewentualnie kiedy) zastąpić tradycyjne badania poligraficzne, stwierdzić trzeba, że służą one generalnie do detekcji kłamstwa identycznie pojmowanego jak w badaniach poligraficznych (świadomej wypowiedzi niezgodnej z prawdą i mającej wprowadzić w błąd). **Zasadnicza różnica między badaniami EEG lub fMRI a tradycyjnymi badaniami poligraficznymi polega – poza tym, że używana jest tu inna aparatura – na tym, że badania te nie polegają**

⁴³ J. Widacki (red.), *Kryminalistyka*, 2. wyd., C.H. Beck, Warszawa 2012, s. 393.

⁴⁴ Por. I. Pavlidis, N.L. Eberhardt, J.A. Levine, *Seeing Trough the Face Deception*, *Nature* 2002, 415, s. 35–35; D.Q.A. Pollina, F. Horvath, J.W. Denver, A.B. Dollins, T.E. Brown, *Development Technologies and Test Formats for Credibility Assessment*, *European Polygraph* 2009, 3, s. 3–4.

⁴⁵ C.E. Obermann, *The effect on the berger rythm of mild affective states*, *Journal of Abnormal and Social Psychology* 1939, 34, s. 84–95.

⁴⁶ M. Dufek, L. Richter, *Soudne psychologické praktikum*, Universita Karlova, Praha 1972, s. 20; M. Dufek, *K problematycie polygrafického vysetrovani v kriminalistice*, [w:] *Doplinkkove studyjne materialy pro kriminalisticky smer pravnickeho studia*, Universita Karlova, Praha 1970, s. 24–25, 55–57, 88–92.

⁴⁷ P.J. Guljajew, J.E. Bychowskij, *Issledovanije emocjonalnowo sostojanija czielowieka w processie proizvodstva sledstwiennowo diejstwija*, *Kriminlistika i Sudiebnaja Ekspertiza*, 1972, 9, s. 108.

⁴⁸ Por. J. Wojciechowski, *Detection of cancealed information with of the p-300 potential amplitude analysis*, *European Polygraph* 2014, 8, 4 (30), s. 167–188.

⁴⁹ Por. J. Vendemia, *fMRI as a method of detection of Deception*, *European Polygraph* 2014, 8, 1(27), s. 5–21.

na wyprowadzaniu wniosków o kłamstwie z wyników analizy fizjologicznych korelatów emocji, które są procesem wtórnym do zmian zachodzących w mózgu, ale na obserwacji i analizie tych ostatnich.

7. Detekcja kłamstwa w oparciu o ocenę zachowania oraz lingwistyczne aspekty wypowiedzi

Psychologowie już od dawna zwracają uwagę na fakt, że wnioskowanie o kłamstwie na podstawie tzw. niewerbalnych wskazówek zachowania obciążone jest dużym błędem⁵⁰. W przypadku obserwacji zachowania nie można mówić o tzw. nosie Pinokia, czyli uniwersalnym dla wszystkich ludzi zachowaniu, które jednoznacznie wskazuje na kłamstwo czy też nieszczerłość wypowiedzi. Jedyne, co można zaobserwować to wskazówki świadczące np. o przeżywaniu konkretnych stanów emocjonalnych, które wiążą się z kłamstwem. Można więc powiedzieć, że w przypadku behawioralnych wskazówek kłamstwa sytuacja jest podobna jak w badaniach poligraficznych: ocenie podlegają korelaty emocji. W przypadku badania poligraficznego – fizjologiczne, obserwowane (i zapisywane) za pośrednictwem aparatu, a więc „instrumentalnie”. W przypadku metod opartych o ocenę zachowania – behawioralne, obserwowane bezpośrednio, bez użycia aparatury.

Zwolennicy metod behawioralnych wskazują na to, że zachowanie jest tą cechą, która podczas komunikacji nie podlega dużej kontroli wolicjonalnej. Osoby starające się wprowadzić w błąd bardziej koncentrują się na tym co mówią (na tym by opowiedzieć kłamstwo w sposób przekonujący i spójny), niż w jaki sposób mówią albo jak się przy tym zachowują. Główną hipotezą leżącą u podstaw oceny zachowania w celu poszukiwania nieszczerłości wypowiedzi jest hipoteza Zuckermana, DePaulo i Rosenthala⁵¹. Zgodnie z nią osoby, które kłamią znajdują się w zupełnie innej sytuacji niż osoby, które mówią prawdę. Po pierwsze, doświadczają dużego obciążenia poznawczego, ponieważ muszą sformułować prawdopodobne kłamstwo, wzbogacić je w szczegóły, uczynić prawdopodobnym i jeszcze zapamiętać je tak, aby ich relacja była spójna w czasie. Po drugie, doświadczają innych emocji niż osoby mówiące prawdę: poczucia winy związanej z okłamywaniem kogoś, strachu, że kłamstwo zostanie ujawnione czy podniecenia związanego z oszukiwaniem. Po trzecie wreszcie, to kłamcy, starają się bardzo kontrolować swoje zachowanie by wywrzeć pozytywne wrażenie na osobie okłamywanej. Na podstawie tej hipotezy sformułowano wniosek, że powyższe cechy sytuacji w jakiej stoi kłamca przekładają się na zachowanie, czyli przez to, że kłamca zwykle dokonuje pewnego wysiłku intelektualnego, czuje czasem poczucie winy czy zdenerwowanie podczas kłamania, różne wskaźniki jego zachowania będą na to wskazywać. Innymi słowy, w przypadku wskaźników behawioralnych nie jest brana pod uwagę bezpośrednio detekcja kłamstwa, tylko detekcja wskaźników świadczących o obciążeniu poznawczym, przeżywaniu konkretnych emocji czy kontroli zachowania. Jak twierdzą niektórzy badacze, w przypadku analiz behawioralnych można ocenić, czy pojedynczy komunikat jest szczerzy czy też nie⁵², choć raczej zwraca się uwagę na to, że analizie podlega całe zeznanie czy też wypowiedź⁵³.

Do tej pory próbowano określić kilkadziesiąt pojedynczych wskazówek kłamstwa (np. wizualne: patrzenie w oczy, obecność gestów ilustrujących czy wokalne: wysokość głosu, pauzy w wypo-

⁵⁰ Np. B.M. DePaulo, J.J. Lindsay, B.E. Malone, L. Muhlenbruck, K. Charlton, H. Cooper, *Cues to deception*, *Psychological Bulletin* 2003, 129 (1), 74–118; S.L. Sporer, B. Schwandt, *Paraverbal indicators of deception: A meta-analytic synthesis*, *Applied Cognitive Psychology* 2006, 20, s. 421–446; V. Hauch, I. Blandon-Gitlin, J. Masip, S. Sporer, *Are computers effective lie detectors? A meta-analysis of linguistic cues to deception*, *Personality and Social Psychology Review*, 2014, s. 1–36.

⁵¹ M. Zuckerman, B.M. DePaulo, R. Rosenthal, *Verbal and nonverbal communication of deception*, [w:] L. Berkowitz (red.), *Advances in experimental social psychology*, 14, s. 1–59, Academic Press, New York 1981.

⁵² Np. na podstawie mikroekspresji, P. Ekman, *Facial expressions*, [w:] C. Blakemore, S. Jennett, (red.), *Oxford Companion to the Body*, Oxford University Press, London 2001.

⁵³ A. Vrij, *Detecting...*, op. cit.

wiedzi). Niestety, opieranie się tylko na tych wskazówkach nie prowadzi do trafnych oszacowań⁵⁴. Badacze zwracają uwagę na to, że być może same wskazówki behawioralne nie są wystarczające do tego, by trafnie diagnozować szczerść wypowiedzi. Zgodnie z Teorią Interpersonalnego Oszustwa⁵⁵ oszukiwanie czy nieuczciwe zeznawanie jest aktem komunikacji i sposób zachowania odbiorcy oszustwa (osoby przesłuchującej) ma bardzo duży wpływ na to, w jaki sposób zostanie ono przedstawione i jak łatwo będzie potem mogło być wykryte. Ponieważ oszukiwanie (nieuczciwe zeznawanie) zwykle jest dynamicznym procesem to odpowiednie kreowanie kontekstu (czyli zaangażowania w interakcję, działania strategiczne, tworzenie odpowiednich wymagań konwersacyjnych itd.) przez odbiorcę oszustwa może ułatwić jego wykrywanie. Jednym z przykładów stworzenia systematycznej oceny zachowania na podstawie wskazówek behawioralnych połączonych z odpowiednim sposobem przesłuchania jest ACID⁵⁶, czy dość kontrowersyjna Behawioralna Analiza Zachowania (BAI)⁵⁷, jednakże jej skuteczność również jest dyskusyjna⁵⁸.

Nieco innym zagadnieniem jest kwestia oceny szczerści zachowania na podstawie analizy lingwistycznej⁵⁹. Podstawą wyróżnienia lingwistycznych wskazówek kłamstwa jest wspomniana wcześniej koncepcja Zuckermana, DePaulo, i Rosenthala. Zwraca się też uwagę na to, że kłamstwo tworzone jest inaczej niż prawdziwa relacja. Kiedy człowiek zeznaje szczerze to korzysta ze swojej pamięci długotrwałej, a najczęściej również z pamięci autobiograficznej. Natomiast kłamca korzystał będzie raczej z pamięci semantycznej oraz specyficznych procesów poznawczych, innych niż osoba mówiąca prawdę⁶⁰. Wszystko to będzie przekładało się na inny sposób formułowania nieuczciwej wypowiedzi u kłamcy np. inny dobór słów, używanie w inny sposób zaimków zwłaszcza wszechogarniających, inny stopień uszczegółowienia wypowiedzi itd. Do tej pory zdiagnozowano ponad 200 lingwistycznych wskazówek, które świadczyć mogłyby bezpośrednio o nieuczciwości zeznania⁶¹. Warto zauważyć, że przy analizie lingwistycznej brane jest pod uwagę całe zeznanie, nie zaś jego fragmenty. Innymi słowy nie można w przypadku analiz lingwistycznych mówić o detekcji kłamstwa rozumianej jako pojedynczy komunikat, a raczej o ocenianiu szczerści całej wypowiedzi (całego zeznania) na podstawie analizy jej werbalnych cech.

8. Detekcja kłamstwa na podstawie analizy treści wypowiedzi

Jeden z kierunków badań nad wykrywaniem kłamstwa/nieuczciwości w zeznaniach dotyczył próby odpowiedzi na pytanie, czy jest możliwe odnalezienie w treści zeznań świadków elementów świadczących o świadomym ich fałszowaniu i jakie to mogą być elementy. Na początku próbowano określać różne kryteria czy sposoby oceny samej treści zeznania (w oderwaniu od zachowania świadka), które wskazywałyby na zaistnienie elementów nieuczciwych, zmyślonych⁶², później zaś dążono do zgrupowaniu tych kryteriów w jednorodną metodę polecającą na systematycznej analizie zeznania w celu odnalezienia tzw. kryteriów nieuczciwości⁶³.

⁵⁴ Por. B.M. DePaulo, J.J. Lindsay, B.E. Malone, L. Muhlenbruck, K. Charlton, H. Cooper, *Cues...*, *op. cit.*

⁵⁵ D.B. Buller, J.K. Burgoon, *Interpersonal Deception Theory*, *Communication Theory* 1996, 6, s. 203–242.

⁵⁶ K. Colwell, C.K. Hiscock-Anisman, A. Memon, L. Taylor, J. Prewett, *Assessment Criteria Indicative of Deception (ACID): An integrated system of investigative interviewing and detecting deception*, *Journal of Investigative Psychology and Offender Profiling* 2007, 4, s. 167–180.

⁵⁷ F.E. Inbau, J.E. Reid, J.P. Buckley, B.C. Jayne, *Criminal interrogation and confessions*, Gaithersburg, MD: Aspen 2001.

⁵⁸ A. Vrij, S. Mann, R.P. Fisher, *An empirical test of the Behaviour Analysis Interview*, *Law and Human Behavior* 2006, 30, s. 329–345.

⁵⁹ Czy też werbalnych wskazówek kłamstwa, por. A. Vrij, *Detecting...*, *op. cit.*

⁶⁰ I. Blandón-Gitlin, E. Fenn, J. Masip, A. Yoo, *Cognitive-load approaches to detect deception: Searching for cognitive mechanisms*, *Trends in Cognitive Sciences* 2014, 18, s. 441–444.

⁶¹ V. Hauch, I. Blandón-Gitlin, J. Masip, S. Sporer, *Are computers...*, *op. cit.*

⁶² Np. F. Arntzen, *Psychologia zeznań świadków*, Państwowe Wydawnictwo Naukowe, Warszawa 1989.

⁶³ M. Steller, G. Kohnken, *Statement analysis: Credibility assessment of children's testimonies in sexual abuse cases*, [w:] D.C. Raskin (red.), *Psychological methods in criminal investigation and evidence*, s. 217–245, New York, NY: Springer 1989.

W przypadku metod opartych o analizę treści wypowiedzi trudno mówić o bezpośredniej detekcji kłamstwa. Metody te bowiem zwykle opierają się na tzw. Hipotezie Undeutch⁶⁴, zgodnie z którą:

- a) zeznania oparte na osobistych doświadczeniach świadka będą różniły się pod względem formy i treści (jakościowo i ilościowo) od zeznań nieszczerých oraz
- b) osoby zeznające szczerze będą różniły się od osób nieszczerých pod względem motywacji, co przekłada się na treść zeznania.

Wskazuje się również na to, że zgodnie z modelem Monitorowania Rzeczywistości (Johnson, Raye, 1981), wspomnienia rzeczywistych wydarzeń różnią się od wyobrażeń i fałszywych wspomnień. Prawdziwe wspomnienia o rzeczywistych, osobistych doświadczeniach odzwierciedlają procesy percepcyjne zachodzące w trakcie ich nabywania (np. zawierają informacje dot. zapachów, smaku, kolorów itp.) oraz emocje, natomiast wspomnienia oparte o wyobrażenia zawierają więcej wnioskowania i procesów poznawczych. Model ten został zaadoptowany przez badaczy zajmujących się oceną zeznań na potrzeby stworzenia narzędzia do wykrywania nieszczerých zeznań (RM – Reality Monitoring, Monitorowanie Rzeczywistości)⁶⁵.

Jak łatwo się zorientować, metody które są stworzone na podstawie ww. hipotez mogą wprowadzić różnicować pomiędzy zeznaniami opartymi na przeżytych faktycznie zdarzeniu a zeznaniami całkowicie zmyślonymi, to jednak nie pozwalają wskazać, które z **elementów** zdarzenia są nieprawdziwe (niezgodne z rzeczywistością) lub nieszczerze (czyli co do których elementów świadek zmyśla). Nie trudno bowiem wyobrazić sobie sytuację, w której świadek wprowadzi szczerze opowiada przebieg np. napadu na bank, ale zataja niektóre informacje (na przykład drobne elementy swojego zachowania) lub nietrafnie przypomina sobie pewne rzeczy (na przykład wygląd sprawców). W takim wypadku używanie metod opartych na analizie treści wypowiedzi może być problematyczne, bowiem dzięki nim będzie można jedynie wskazać, że świadek faktycznie przeżył napad na bank, natomiast bardzo trudno będzie wskazać które elementy jego wypowiedzi są celowo sfalszowane. Innymi słowy, podobnie jak w przypadku metod lingwistycznych czy niektórych metod opartych na analizie zachowania) przedmiotem zainteresowania w metodach opartych na analizie wypowiedzi nie jest detekcja pojedynczego kłamstwa co do jakiegoś pojedynczego szczegółu, czy jak w badaniu poligraficznym do przestrzeni ograniczonej pytaniami testów, tylko ocena szczerości świadka rozumianej jako relacjonowanie o zdarzeniu, które zostało faktycznie przeżyte.

Dlatego też początkowo wykorzystywano takie metody jak np. CBCA (*Criteria Content – Based Analysis*, oparta na kryteriach ocena zeznania) tylko do oceny zeznań dzieci molestowanych seksualnie⁶⁶, a obecnie można stwierdzić że metody te są użyteczne również przy zeznaniach dzieci – ofiar innych typów przestępstw⁶⁷, dzieci będących świadkami⁶⁸, dorosłych świadków⁶⁹ jak i dorosłych ofiar⁷⁰ oraz sprawców przestępstw⁷¹.

⁶⁴ U. Undeutsch, *Beurteilung der Glaubwürdigkeit von Zeugenaussagen*, [w:] U. Undeutsch (red.), *Handbuch der Psychologie*, Band 11: *Forensische Psychologie*, Göttingen: Hogrefe 1967.

⁶⁵ S.L. Sporer, *Reality monitoring and detection of deception*, [w:] P.A. Granhag, L.A. Strömwall (red.), *Deception of detection in forensic contexts*, Cambridge University Press, Cambridge 2004, s. 64–102.

⁶⁶ Głównie w celu stwierdzenia, czy molestowanie faktycznie miało miejsce – np. A. Craig, R. Scheibe, D.C. Raskin, J.C. Kircher, D. Dodd, *Interviewer questions and content analysis of children's statements of sexual abuse*, *Applied Developmental Science* 1999, 3, s. 77–85.

⁶⁷ K. Pezdek, A. Morrow, I. Blandon-Gitlin, G.S. Goodman, J.A. Quas, K.J. Saywitz, *Detecting deception in children: Event familiarity affects Criterion-Based Content Analysis ratings*, *Journal of Applied Psychology* 2004, 89, s. 119–126.

⁶⁸ A. Vrij, L. Akehurst, S. Soukara, R. Bull, *Let me inform you how to tell a convincing story: CBCA and Reality Monitoring Scores as a function of age, coaching and deception*, *Canadian Journal of Behavioural Science* 2004, 36, s. 113–126.

⁶⁹ *Ibidem*.

⁷⁰ K. Landry, J. Brigham, *The effect of training in Criteria-Based Content Analysis on the ability to detect deception in adults*, *Law and Human Behavior* 1992, nr 16, s. 663–675.

⁷¹ S. Porter, J.C. Yuille, D.R. Lehman, *The nature of real, implanted, and fabricated memories for emotional childhood events: Implications for the recovered memory debate*, *Law and Human Behavior* 1999, 2, s. 517–537.

Reasumując, w przypadku metod opartych na analizie treści zeznania trudno mówić o „detekcji kłamstwa” (tak jak w przypadku badań poligraficznych) – bardziej trafne wydaje się określenie „ocena szczerości zeznania” czy jeszcze ściślej „ocena tego, jak bardzo prawdopodobne jest, że świadek przeżył w ogólnych zarysach opisywane wydarzenie”.

Jedną z najczęściej używanych przez biegłych w USA, Niemczech czy Holandii metodą oceny szczerości treści zeznania jest procedura SVA (Ocena jakości zeznania – *Statement Validity Analysis*)⁷². Procedura ta składa się z czterech części. W pierwszej biegły zapoznaje się z aktami sprawy, głównie w celu poszukania danych o świadku i zdarzeniu oraz przeanalizowania dotychczas złożonych zeznań. W części drugiej następuje ustrukturyzowane przesłuchanie świadka, które powinno być nagrane, a następnie przepisane, ponieważ następne części procedury dotyczą analizy tylko i wyłącznie treści wypowiedzi. Analiza prowadzona np. na materiale wideo jest błędna, ponieważ na ocenę biegłego może mieć wpływ zachowanie świadka, jego wygląd czy inne czynniki pozamerytoryczne. Również nieprawidłowe jest prowadzenie analiz treści zeznania na protokołach z przesłuchania, bowiem na ogół nie zawierają one dokładnych wypowiedzi świadka, a jedynie przeformułowane i opracowane przez osobę przesłuchującą treści. Nadto zakres wypowiedzi osoby przesłuchiwanej, utrwalanej w formie protokołu, stymulowany jest często pytaniami osoby przesłuchującej, które nie są odnotowane w protokole.

Trzecim elementem procedury SVA jest Oparta na Kryteriach Analiza Treści (wspomniane wcześniej CBCA) a ostatnim tzw. Lista Kontrolna Prawdziwości (*Validity Checklist*), w ramach której sprawdza się w uzyskanym zeznaniu cechy psychologiczne (np. czy świadek posługuje się językiem stosownym do jego wieku i wiedzy), cechy samego przesłuchania (np. czy nie było prowadzone w sposób sugerujący), aspekty motywacyjne (np. dlaczego świadek zdecydował się złożyć zeznanie) oraz zgodność z innymi zeznaniami czy prawami natury.

Zgodnie ze słowami twórców⁷³ najważniejszym krokiem w SVA jest przeprowadzenie przesłuchania a następnie CBCA. Samo CBCA składa się z 19 kryteriów dotyczących semantycznych cech treści zeznania, które częściej pojawiają się w zeznaniach szczerych (dotyczących zdarzenia, którego osoba rzeczywiście była świadkiem) niż nieszczerych (dotyczących zmyślonych zdarzeń). Te kryteria to np. logiczna struktura wypowiedzi (kryterium 1), duża ilość szczegółów (kryterium 3) czy odtwarzanie wypowiedzi (kryterium 5). Zakłada się, że w zeznaniach prawdziwych kryteria te będą występować w większym nasileniu, niż w zeznaniach fałszywych. Mówiąc dokładniej, w zeznaniach szczerych ludzie przeszkoleni w używaniu metody CBCA powinni częściej stwierdzać istnienie logicznej struktury wypowiedzi, diagnozować większą ilość szczegółów czy częściej stwierdzać istnienie wypowiedzi lub części rozmów są przytaczanych w ich oryginalnej formie niż w zeznaniach zmyślonych. 19 kryteriów prawdziwości zostało określonych na podstawie wielu lat doświadczeń i obserwacji, a następnie zostały zweryfikowane w licznych badaniach naukowych. Ogólnie można powiedzieć, że większość kryteriów faktycznie pojawia się częściej i w większym natężeniu w zeznaniach szczerych, tj. dotyczących faktycznie przeżytych zdarzeń⁷⁴.

Podobną, choć nie tak kompleksową metodą oceny treści zeznania jest Monitorowanie Rzeczywistości (RM). Narzędzie to składa się z 8 lub 7 kryteriów⁷⁵. Kryteria te są oparte na dobrze opisanych i zweryfikowanych teoriach psychologicznych dotyczących formowania się wspomnień i dotyczą np. informacji przestrzennych (w zeznaniach szczerych powinno ich być więcej), przejrzystości (zeznania szczere powinny być bardziej żywe i przejrzyste) czy uczuć (im więcej emocji opisanych w zeznaniu tym większe prawdopodobieństwo, że jest oparte na faktycznie przeżytym zdarzeniu). Jakkolwiek może wydawać się, że metody takie jak SVA czy RM są dość łatwe do nauczenia się, to badacze

⁷² M. Steller, G. Kohnken, *Statement...*, *op cit.*

⁷³ G. Kohnken, *Statement Validity Analysis and the detection of the truth*, [w:] P.A. Granhag, L.A. Strömwall (red.), *The detection of deception in forensic contexts*, Cambridge University Press, Cambridge 2004, s. 41–63.

⁷⁴ R. Volbert, M. Steller, *Is this testimony truthful, fabricated, or based on false memory? Credibility assessment 25 years after Steller and Köhnken (1989)*, *European Psychologist* 2014, s. 207–220.

⁷⁵ S. L. Sporer, *Reality...*, *op cit.*

wskazują, że aby umieć dobrze odszukiwać kryteria szczerości w zeznaniach potrzebny jest długo-trwały trening⁷⁶.

Ocenianie zeznania przy użyciu CBCA czy RM mniej więcej wygląda w ten sam sposób i dzieje się zwykle w dwóch fazach. W fazie pierwszej biegły poszukuje kryteriów prawdziwości w danym zeznaniu, a następnie przypisuje im wartość punktową. Zwykle jest to 0 – jeżeli dane kryterium nie występuje, 1 – jeżeli występuje w słabym nasileniu lub 2 – jeżeli występuje w sposób bardzo wyrazisty. Następnie, w przypadku badań eksperymentalnych, sumuje się punkty uzyskane przez zeznania i przy pomocy statystycznych analiz porównuje wyniki uzyskane przez zeznania szczere i nieszczere i ustala, czy istnieją istotne różnice między nimi. Trzeba jednak podkreślić, że w praktyce ta metoda nie ma zastosowania – po prostu biegły ma do dyspozycji tylko jedno zeznanie co do którego musi ocenić, czy jest szczere czy też nie. Ponieważ dla żadnej z ww. opisanych metod analizy treściowej nie ma obiektywnie ustawionych granic, od ilu punktów zaczyna się szczere zeznanie, biegły na podstawie znajomości narzędzia, poprawnie wykonanej analizy, wiedzy psychologicznej i innych czynników musi sam podjąć decyzję co do oceny szczerości danego zeznania.

Warto tutaj zauważyć, że problem ten pojawia się również przy metodach instrumentalnych (jak poligraf czy termografia) i innych metodach nieinstrumentalnych. W przypadku każdej z tych technik osoba dokonująca analizy szczerości/prawdziwości relacji na podstawie wskaźników metody (kryteriów prawdziwości w przypadku CBCA, kryteriów szczerego zachowania werbalnego w przypadku analizy zachowania, danych z trzech kanałów poligrafu itd. musi ocenić, czy zeznanie jest prawdziwe czy też nie. Oznacza to, że **nawet jeżeli wskaźniki danego narzędzia są dobrze przebadane i dość dobrze wskazują na prawdziwość relacji, to już wnioski do jakich dochodzą osoby używające danych metod mogą być błędne. Przyczyną takiego stanu rzeczy może być najprawdopodobniej fakt, że specjaliści używający metod określania szczerości zeznań czy wykrywania kłamstwa dokonując ostatecznego rozstrzygnięcia, czy zeznanie jest prawdziwe czy fałszywe nie zawsze kierują się wskazaniem danej metody, tylko własnymi przekonaniem i pozamerytorycznymi przesłankami.**

9. Ocena tzw. wiarygodności psychologicznej

Wiarygodność psychologiczna nie jest wprawdzie metodą detekcji czy szacowania prawdziwości zeznań, jednakże jest zagadnieniem które bezpośrednio z taki metodami wiąże. W praktyce dość często zdarza się, że biegli psychologowie wypowiadają się w kwestii „wiarygodności” czy „prawdomówności” świadka lub niekiedy nawet oskarżonego. Zdarza się, że wręcz są do tego zachęceni przez organ procesowy. Pomijając już wzgląd formalny, że ocena wiarygodności (czy prawdomówności) świadka jest oceną dowodu i jako taka jest zastrzeżona dla organu procesowego i w tym zakresie biegły nie może organu procesowego wyręczać, trzeba stwierdzić wyraźnie, że psycholog nie ma narzędzi do dokonania takiej oceny. **Tak więc ilekroć biegły psycholog wypowiada się na temat wiarygodności świadka, po pierwsze narusza procedurę karną, poza tym wypowiada się nie w oparciu o wiedzę psychologiczną, ale na podstawie swej wiedzy potocznej i doświadczenia życiowego, a więc działa poza granicą „wiadomości specjalnych” o których mowa w art.193 kpk⁷⁷.** Biegły psycholog może jedynie ocenić aprioryczną wiarygodność świadka, uwzględniając takie elementy jak cechy osobowości (np. podatność na sugestie czy skłonność do manipulowania ludźmi), zdolność do spostrzegania i odtwarzania spostrzeżeń, ewentualną skłonność do konfabulacji (czyli wypełniania luk pamięciowych treściami zmyśloną) itd.

Wszystkie te cechy mogą ułatwiać składanie zeznań subiektywnie prawdziwych, ale obiektywnie fałszywych. Tak więc ich stwierdzenie może być ważną wskazówką ułatwiającą organowi procesowemu ocenę takiego dowodu w szczególności sugerującą mu ostrożność w takiej ocenie. **Niestwier-**

⁷⁶ V. Hauch, S.L. Sporer, S.W. Michael, Ch. A. Meissner, *Does training improve the detection of deception? A meta-analysis*, Communication Research 2014.

⁷⁷ Por. J. Widacki, *Próby weryfikowania prawdomówności w procesie karnym*, Palestra 2012, 3–4, s. 13–18.

dzenie takich cech nazywane bywa czasem stwierdzeniem „wiarygodności psychologicznej”⁷⁸. Ta „wiarygodność psychologiczna” jest oczywiście czymś zupełnie innym od wiarygodności konkretnej wypowiedzi, konkretnego świadka. Mówiąc potocznie, wiarygodność psychologiczna świadka (jako źródła informacji) ogranicza się do stwierdzenia, że dany świadek, jak nie chce to nie kłamie. **Czyli świadek „wiarygodny psychologicznie”, to świadek u którego nie stwierdzono zaburzeń spostrzegania, zapamiętywania, odtwarzania spostrzeżeń, ani takich cech osobowości, które sprzyjać mogą składaniu zeznań obiektywnie fałszywych, choć subiektywnie prawdziwych.**

Ale w praktyce spotkać można inne rozumienie „wiarygodności psychologicznej”, którą biegły w sposób nieuprawniony odnosi nie do źródła, osoby świadka, ale do jego zeznania. Tak np. w jednej ze spraw biegły psycholog napisał „z psychologicznego punktu widzenia, treść zeznania małoletniego J.G. posiada znamiona wiarygodności”⁷⁹. Jak wynika z ustnej uzupełniającej opinii, podstawą takiego twierdzenia biegłego było ustalenie badaniem psychologicznym „braku podstaw do stwierdzenia u badanej tendencji do konfabulacji lub zaburzeń pamięci”.

Jak jeszcze raz warto podkreślić, posiadanie przez świadka odpowiednich predyspozycji do składania zeznań (prawidłowo rozwinięta osobowość czy zdolność do poprawnego odtwarzania spostrzeżeń) nie determinuje tego, że jego zeznania będą szczere. Świadek może bowiem mieć intencję wprowadzenia wymiaru sprawiedliwości w błąd i zeznawać nieszczerze. Dlatego biegli psychologowie używają czasem terminu **„wiarygodność psychologiczna zeznania”** dla określenia, czy złożone zeznanie nosi tzw. psychologiczne znamiona wiarygodności. Najczęściej są one rozumiane czy utożsamiane z opisanymi wcześniej treściowymi kryteriami szczerości (głównie zaczerpniętymi z CBCA) lub opisane jako wynikające z procesu formowania zeznań (SVA). Innymi słowy, biegły opisując psychologiczną wiarygodność zeznania **odnosi się nie do cech świadka, tylko cech konkretnego zeznania, złożonego w konkretnej sprawie** – a dokładniej tego, czy można w nim odnaleźć kryteria szczerości. Pamiętać jednak trzeba, że wartość diagnostyczna sformalizowanych procedur takich jak SVA czy CBCA jest daleka od 100%. **Zatem posłużenie się niewystandaryzowaną metodą, parafrazą CBCA czy SVA może dać wynik jedynie ogólnie informujący, orientacyjny, nigdy zaś rozstrzygający. Jest bardziej wynikiem osobistego przekonania biegłego, niż rezultatem jakiegoś pomiaru.**

Jednak nawet taki orientacyjny wynik analizy tzw. „psychologicznej wiarygodności zeznań świadka” czyli ocena procesu formowania zeznania (w jaki sposób i w jakich okolicznościach było składane), zmiennych wpływających na formowanie się zeznania oraz w konsekwencji – ewentualnych kryteriach wskazujących na szczerość/nieszczerość danego zeznania może być pomocna dla sędziego doświadczonego i rozumiejącego ograniczenia metody i ułatwiać mu ocenę wartości dowodowej składanego zeznania.

Podsumowanie

Jak trzeba podkreślić, w kryminalistyce i psychologii sądowej nazwy „kłamstwo” i „detekcja kłamstwa” funkcjonują w różnych znaczeniach, różny jest też przedmiot „detekcji kłamstwa”. Poniższa tabela pokazuje zestawienie funkcjonujących obecnie metod detekcji wraz z krótką charakterystyką założeń.

Jak wynika z tabeli 1, obecnie istniejące narzędzia określane często jako służące do wykrywania kłamstwa, w istocie bezpośrednio wcale go nie mierzą. W większości przypadków zakłada się, że (1) kłamstwo powoduje zmiany emocjonalne, poznawcze itp. i że (2) te zmiany emocjonalne i poznawcze przekładają się na obserwowalne cechy (zmiana fizjologiczna, zmiana zachowania itd.) i w rezultacie na podstawie zmian cech obserwowalnych można przeprowadzić wnioskowanie o co najwyżej o zaistnieniu nieszczerości wypowiedzi czy całych zeznań. W przypadku, kiedy jedno z założeń jest niespełnione (np. osoba kłamiąca nie odczuwa żadnych emocji z tym związanych) używanie takich narzędzi przestaje mieć sens. Stąd też wielki entuzjazm związany z ideą badania kłamstwa przy użyciu fMRI. W założeniu zwolenników tej metody użycie jej prowadziłoby do zaobserwowania bezpośred-

⁷⁸ *Ibidem*, s. 14.

⁷⁹ Sprawa 2 Ds. 1386/13 Prokuratury Rejonowej w K. N.H.

nio samego faktu tworzenia się kłamstwa w miejscu, gdzie ono powstaje (np. określonych strukturach mózgu)⁸⁰ nie zaś wnioskowania o kłamstwie na podstawie obserwowania jego zakładanych korelatów. Dodatkowo używanie fMRI wg jego zwolenników pozwoliłoby na oszacowanie, czy poszczególne bardzo proste komunikaty („Nie zabiłem go”) czy nawet słowa są wypowiedziane nieszczerze. Należy jednak podkreślić, że ta metoda przy obecnym stanie wiedzy jest jeszcze bardzo zawodna.

Tabela 1. Zestawienie metod detekcji kłamstwa i badanych przez nie korelatów

	Badania poligraficzne techniką IR	Badania poligraficzne techniką CQ i GKT (CIT)	Inne psychofizjologiczne (analiza głosu, zmiany temperatury twarzy)	Techniki oparte o ocenę pracy mózgu (EEG, fMRI)	Techniki oparte o analizę zachowania niewerbalnego	Techniki oparte o analizę treści wypowiedzi (np. CBCA, RM, SCAN oraz o wskazówki lingwistyczne)
Zakładany korelat kłamstwa	Zmiany emocjonalne	Zmiany emocjonalne	Zmiany emocjonalne	-	Zmiany emocjonalne i poznawcze	Procesy poznawcze i motywacyjne
Obserwowalna (badana) cecha	Zmiany fizjologiczne	Zmiany fizjologiczne	Zmiany fizjologiczne	Zmiany funkcjonowania mózgu	Zmiany zachowania niewerbalnego	Treść wypowiedzi
Wnioski	Ocena szczerości w obszarze wyznaczonym sumą pytań krytycznych	Ocena szczerości odpowiedzi na poszczególne pytania krytyczne	Ocena szczerości w obszarze wyznaczonym sumą pytań krytycznych	Ocena szczerości poszczególnych odpowiedzi	Ocena szczerości wypowiedzi	Ocena, czy zeznanie jest relacją przeżytych zdarzeń

Warto zatem jeszcze raz podkreślić, że obecnie używane w praktyce metody wykrywania kłamstwa w swojej istocie badają jedynie obserwowalne i założone korelaty kłamstwa oraz pozwalają jedynie na ostrożne wnioskowanie o szczerości całych, szczegółowo określonych wypowiedzi, przy czym zdaje się nie ulegać wątpliwości, że wartość diagnostyczna badania poligraficznego jest z wszystkich omawianych metod zdecydowanie najwyższa, porównywalna z wartością diagnostyczną innych rutynowo stosowanych w kryminalistyce metod identyfikacji, wystarczającą dla każdego dowodu pośredniego.

Streszczenie

Artykuł poświęcony jest zagadnieniu metod „wykrywania kłamstwa” w kryminalistyce i psychologii sądowej. Szczegółowo opisane są kwestie filozoficzne, językowe i psychologiczne związane z pojęciem kłamstwa. Przybliżone są metody potocznie określone jako „służące detekcji kłamstwa” wraz z dokładną analizą przedmiotu ich detekcji oraz zakresu użyteczności w praktyce. W artykule poruszone są również kwestie wiarygodności psychologicznej świadka i zeznania.

⁸⁰ Trzeba zaznaczyć, że takie stwierdzenie jest daleko idącym uproszczeniem, por. E. Rusconi, T. Mitchener-Nissen, *Prospects of functional magnetic resonance imaging as lie detector*, *Frontiers in Human Neuroscience* 2013, 7, s. 1–12. por. również: J. Vendemia, *fMRI as a method of detection of deception: a review of experience*, *European Polygraph* 2014, 8, 1 (27), s. 5–21.

Jan Widacki

prof., Department of Forensics, Criminology and Police Science
Andrzej Frycz Modrzewski Krakow University

Karolina Dukąła

doctoral student at the Institute of Psychology at the Jagiellonian University

Lie detection – of what?*

The article discusses the issue of methods of “lie detection” in criminology and forensic psychology. Considerations are described in detail philosophical, linguistic and psychological associated with the concept lies. Approximate methods are commonly referred to as “used to detect lies” along with a thorough analysis of the subject of their detection and usability in practice. The article raised issues of credibility are also psychological and witness testimony.

Key words: deceit, lie detection, psychological credibility of the witness and testimony

Introduction

Classic criminologist Hans Gross claimed that “most of the work of forensics is a struggle against lies”¹. No wonder that in the literature of criminology and forensic psychology, the name “lie detection”, “instrumental detection of lies”, “non-instrumental detection of lies” (*lie-detection*)² is often used. The literature also includes such terms as “fraud detection”³, or more precisely, “deception” (*detection of deception*) and “a psychological evaluation of the reliability of testimony”⁴. Apart from this, forensic psychology uses the term “psychological credibility”, which is a criterion for separating the sincere from the false.

Already in the nineteenth century, Pierre Simone de Laplace proposed dividing testimony into different categories. He distinguished true testimony, that is that which reflect reality, and sincere testimony, which are an accurate reflection of the subjective reality of the witness (i.e. the witness has intention to testify truly, according to what he remembers), but cannot reflect the objectively existing reality. These criteria allow for the creation of four types of categories of testimony:

- Genuine and sincere (reflecting objective reality and subjectively true);

* Jan Widacki, Karolina Dukąła, *Lie detection – of what?*, “Problemy kryminalistyki” 2015, 287 (1), pp. 61–74.

¹ H. Gross, *Criminal psychology: A manual for judges, practitioners and students*, Little, Brown and Comp., Boston 1918, p. 474.

² ct. e.g. C.D. Lee, *The instrumental detection of deception*, Ch. Thomas, Springfield 1953; A. Vrij, *Detecting Lies and deceit. The psychology of lying and implications for professional practice*, John Wiley & Sons, Chichester, 2000; J. Ulatowska, *Knowledge of cues to deception – looking for its determinants*, “Problems of Forensic Sciences” 2009, 80, pp. 411–428; *Kryminalistyka*, J. Widacki (red.) C.H. Beck, Warsaw 2012, pp. 106–109, pp. 393–395.

³ K. Cantarero, *Detecting deceit in interpersonal communication*, “Social Psychology” 2009, 4, pp. 167–176.

⁴ B. Wojciechowski, *Psychological analysis of the content of the testimony of witnesses and evaluation of their credibility*, “Palestra” 2012, No. 1–2, pp. 70–80.

- False and insincere (witness seeks to ensure that convey a false data that are inconsistent with the objective truth);
- False, but sincere (witness tries to testify in accordance with its knowledge and is convinced that telling the truth, but his statement, e.g. As a result of errors of memory, forgetfulness etc. untruthful objective) and
- Insincere, but real (very rare group of evidence, the witness wants to lie and testify accordance with what is remembered, but still provides information in accordance with objective truth).

Most current methods of “lie detection” therefore refer to a category of subjectively insincere and objectively false. It is possible to meet the claim that the object of the detection method is not a lie or sincerity but **“guilty knowledge”** (*guilty knowledge*)⁵ or simply “guilty” (*guilty*)⁶ “knowledge about the realities of the act”⁷ or **“emotional relationship of the examinee with the act”** or **“concealed information”** (*concealed information*)⁸. However, whenever there is talk about “guilty knowledge” or “knowledge of the realities of the act” or an “emotional connection of the examinee with the act” there will always be a hidden (withheld) guilty knowledge, withheld knowledge of the act, concealment of an emotional relationship with the act.

Thus, at first glance, the term “lie” as the object of detection or evaluation is used in forensic and psychological literature in various senses, in addition meanings not entirely certain, so tidying up the terminology seems necessary.

Issues of language

According to the dictionary of contemporary Polish language, “a lie” is a claim inconsistent with reality, uttered with the intention of bringing someone into error⁹. But according to the dictionary, “lie” is also “false”, “untrue”. In the dictionary of the Polish language from 1961¹⁰ added still one meaning of the word “lie”, and more specifically, another synonym: “lying”. This dictionary also introduces something important, as it turns out, for further consideration, **the distinction between “negative lie” and “positive lie”**¹¹. **With the first, named “negative lie”, we are dealing, for example, when a person asked by someone else, does not admit to what he did. With “positive lie” we are dealing with when someone fabricates something which indeed did not take place.** In turn, modern dictionaries of Polish language for the word “lie” present several dozen synonyms, grouping them in more than 40 major groups of meaning.

It seems there is no doubt **that the essence of a lie is a statement that forwards information knowingly untrue.** False information may take any form of communication. This may take the form of sentences uttered verbally or in writing. It can also assume the form of a sign (facial expressions, pantomime) or gesture. In the case of forwarding information in an untrue statement (spoken or written), it is logically false. In addition, the speaker (or writer) is aware or, at least, is not sure of logical value of the spoken (or written) statement, but forward them as if he was sure he was passing a true statement.

⁵ D.T. Lykken, *The validity of the guilty knowledge technique*, “Journal of Applied Psychology” 1960, 44, pp. 258–262; D.T. Lykken, *Guilty Knowledge Test – the right way to use lie-detector*, “Psychology Today” 1975, 8, 10, pp. 56–60.

⁶ D. T. Lykken, *The GSR In detection of guilt*, “Journal of Applied Psychology” 1959, 43, pp. 385–388.

⁷ M. Kulicki, *Criminology. Selected problems of the theory and practice of forensic investigation*, Toruń 1994, p. 482.

⁸ Por. M. Nakayama, *Practical use of the Concealed Information Test for criminal investigation in Japan* [in:] M. Kleiner, *Handbook of polygraph testing* (ed.), Academic Press, London, 2002, pp. 49–86.

⁹ *Concise Polish dictionary*, ed. E. Sobol, Ed. PWN, Warsaw 1996, p. 334.

¹⁰ W. Doroszewski (ed.), *Dictionary of Polish*, Polish Academy of Sciences, Warsaw 1961, vol. III, lit. H–K, p. 746.

¹¹ *Ibidem*, p. 746.

Thus, a prerequisite for the existence of a lie is a conscious declaration of a false statement.

Is it also a sufficient condition? Is every false statement a lie? Or maybe just one that is addressed to someone, moreover, spoken in order to lead the recipient into error?

In other words, is Smith lying if he says into a mirror, while standing in an empty room, making the untrue statement, “I am Napoleon Bonaparte”? And what about the case where Smith, at dinner, says to his wife, who has known him for decades, and, in addition, is aware of history, “I am Napoleon Bonaparte”? In both cases, Smith makes a false statement. Do we consider it a lie, or rather a joke, prank? Is it proper to consider Smith a liar, or rather a prankster, or a buffoon?

Another problem that emerges here is the question of whether an intentionally misleading gesture, which is by definition not a statement, even more (in a logical sense) a lie? When someone asks a secretary if the director is in his office, and he is, and the secretary responds to the negative by a shake of the head, misleading the questioner; is it a lie or not a lie?

If we consider that she lied, that stems from these two important conclusions:

1. A lie does not require termination of a full false statement;
2. The essence of lying is a desire to mislead the recipient of a statement, but not the form in which the message is conveyed.

Gesture (nodding the head, contradicting move of the head or the hand, helpless spreading of hands and the like) are replacing words in the process of the communication: “yes”, “no”, “I do not know”. Can such a gesture or one-word utterance (“yes”, “no”) can be logically true or false? That mere utterance of course not, but in the context of questions may have logical value.

If to the question, “Is the director is in his study?”, the secretary replies with the words “yes” or “no” or a gesture replacing these words, a message was actually passed stating either “the director is in his office” or “the director is not in his office”, and these sentences are boolean.

Philosophical Issues

From the previous considerations, **lying is to give the recipient a message that is false with the intention that the recipient recognizes it as true.**

For philosophy, lying is primarily an ethical issue. Is a lie (*lie*) a specific evil that man introduced to nature¹² or does it occur in nature, for example, in the form of mimicry (assimilation into the environment or other species). Is mimicry (the assimilation of organisms vulnerable to organisms capable of defence), common in animals, and even plants, or a more sophisticated behaviour of animals designed to deceive enemy or prey. If these behaviors are considered a form of lie, human deceit would stand out in one form (language), sometimes with a refined purpose and degree of awareness.

A justified lie is another ethical problem. Ethical review is associated not so much with a lie but rather with the purpose for which it was used. A lie can be ethically justified in medicine¹³ when, for example, one does not want to deprive the patient of hope and conceal from him the actual state of his health. A lie can be justified also in the politics, when during an interview, a foreign state is being provided with disinformation, when an intelligence agent is pretending that he is a diplomat. A lie can be justified in everyday life when we are tell the unmerited, but also an expected compliment to somebody. An ethically justified lie may be when raising children, when, for example, you do not tell a small child that his father died or landed in prison, but only saying – of course a false statement – that the father left. To all intents and purposes we are also lying telling fairy tales to a child or saying that a gift was brought by St. Nicholas etc. A lie can also be justifiable in business if, for example, in

¹² M. Buber, *Good and Evil*, Charles Scribner's Son, New York, p. 7.

¹³ A. Szczekliki, *Truth for a seriously ill*, “Diaretos to” 2005, 4, 151; E. Kubler-Ross, *Talks about death and dying*, Media, Family, Poznan 1998, pp. 55–65; cf. also J. Kucharski, *Justified lie in contemporary applied ethics*, University of Ignatianum – WAM Publishing House, Cracow 2014, p. 227 et seq.

the course of negotiations we use a bluff. A recently issued monograph by J. Kucharski¹⁴ is devoted to the justified lie.

It seems that a worse problem than a lie, recognized by Dupart¹⁵ as “the most terrible social evil”, is the prospect of the complete elimination of lies, a topic of brain research, possibly not so far in the future, is the possibility of reading human thoughts, which would not only eliminate the lie, but would end our civilization in its present term¹⁶. It would deprive us not only of any privacy, but also force us to reveal the truth, even where the lie is ethically justified and necessary.

For philosophy, the lie is not only an ethical issue, but also epistemological. In an attempt to define the lie, like it or not, we come to the notion of the truth. The notion of the truth is under consideration both in logic (especially logical semantics) and philosophy, which is the central problem of epistemology.

From the philosophical account, truth is a feature of our cognition, rather than a property of reality. In other words, the truth refers not to fact, only to the judgement of fact, to statements made about the facts.

Western philosophy tries to determine what is true, how to find it and according to what criteria. Proposals are many. One of them, historically the oldest but still current, is the classic concept of truth. According to this proposal (sometimes called a theory), the truth is the conformity of things and mind (*veritas est adequatio rei et intellectus*), or the compatibility of reality and judgement on this reality. Aristotle, who is considered the precursor of this concept, wrote in *Metaphysics* that “To say of what is that it is not, or of what is not that it is, is false, while to say of what is that it is, or of what is not that it is not, is true”¹⁷.

At first glance, everything here is clear and consistent with intuition. The problem begins when we want to determine, on what that “agreement” (*adequatio*) of the thing and the judgement of the thing relies. In particular, is the real state of affairs similar to the thought about it? Whether it is here about a literal accuracy, whether for the belief that this accuracy exists? The difficulties posed by the classical concept of truth try to break the so-called non-classical concepts of truth. One of them is the concept of obviousness, also called the concept of evidence. According to it, the truth is the judgement of reality that is obvious. Neo-Kantian Heinrich Rickert (1863–1936) believed that obvious proposition imposes the necessity, felt as a compulsion or obligation, to accept it¹⁸. Another concept is the concept of non-classical universal agreement. Yet another non-classical concept is the concept of coherence, according to which the truth of properly constructed sentences (judgement) depend on its compliance with other sentences, already recognized as true. It is therefore a logical compatibility consisting of non-contradiction within the framework of a system.

Another concept is the utilitarian (pragmatic) concept, according to which the truth lies in the fact that its adoption leads to useful, practical goals. If recognition of a true statement (judgement) or a theory is useful, practical, this statement (theory) is true.

The lie in terms of the psychological

The psychology of the lie is defined in different ways. In developmental psychology, there are three approaches to lies: The **propositional** (a lie is when one person passes false information to another in order to place them in error)¹⁹; **prototypical** (any information can be put on a continuum, with at one

¹⁴ J. Kucharski, op.cit.; ct. also: J. Widacki, A. Szuba-Boroń: Book Review: J. Kucharski, op.cit., “European Polygraph”, 2014, 4, pp/ 199–203.

¹⁵ G.L. Dupart, *Lie. Psychosociological study*, Gebethner and Wolff, Warsaw 1905, pp. 9–10.

¹⁶ J. Widacki, *Farewell to the present day – science and technology against terrorism. Terror science and technology* [in] G. Babinski, M. Kapiszewska (eds.), *Understanding the Present. Jubilee book Prof. H. Kubiak*, AFM Publishing House, Cracow 2009, pp. 287–294.

¹⁷ A. Węgrzecki, *Outline of philosophy*, AE publishing company, Cracow, 2005, p. 108.

¹⁸ Ibidem, p. 110.

¹⁹ R.M. Chisholm, T.D. Feehan, *The intent to deceive*, “Journal of Philosophy” 1977, 74, pp. 143–159.

end the prototypical lie and the second the prototypical truth: the real message that the sender transmits to another person with the intention of informing them²⁰ and the so-called **folkloristic model of lies**²¹ where it is assumed that a lie is just such a message, which is intended to harm the other person. Thus the altruistic lie (for example, undeserved compliments to another person in order to please them) is not a lie, according to this assumption.

In the psychology of communication, it is assumed that a lie is a kind of communication given by the sender to the recipient and the result of interaction between partners of this interaction²². In other words, a lie can be divided into all aspects of the communication process: sender, recipient, code, context, transmitted information (the lie itself) and feedback. As Cantarero indicated²³, research of the lie can concern all these aspects. And so, research on the context relate to environmental influences, including other people in the room to the frequency or quality of lying, the study of the sender and receiver are focused on the determination of personality traits, temperament or strategies used as well as the deceived persons, who were influenced by the efficiency of the lie. Research on codes (e.g. words, gestures, movement) and messages in turn relate to methods by which lies can be detected.

Psychologists dealing with lying in a social or judicial context usually determine them as a false communication, of which benefitting the lying person is the purpose²⁴, or more precisely, an act of which has the purpose of producing a belief in another person, as for which liar is convinced that it is false²⁵. Thus, for example, false memories (on say, harassment, which did not take place) are not considered as deceit (a lie).

The lie as a matter of research in criminology and forensic psychology

Both the positive lie and the negative lie are subjects of interest in criminology and forensic psychology. A negative lie consists in the denial of reality, for example, when a suspect does not admit to committing an offence (which they committed). We are dealing with the positive lie when, e.g., somebody untruthfully accuses someone. You could easily imagine a situation when we have to deal simultaneously with a negative and positive lie, e.g., someone not admitting to an act committed by them, who lays the blame on someone else. **In practice we are interested only in situations when someone consciously makes false statements with the intention that somebody else will regard them as true.** It may be both statements that in a logical sense could mean this or that, which only in the context of a question becomes a statement making logical sense and a false statement but of single words (yes-no) or replaced with gestures. This raises the question of whether we have today techniques (instrumental or noninstrumental) allowing for the detection of lies understood this way. It is difficult to give a clear answer to this question. **So far, attempting to detect a lie (meeting the conditions listed above) makes sense only when uncovering the lie comes with some sort of consequences for the liar.** This is because existing methods only allow lie detection to detect physiological and cognitive correlates of emotions that have accompanied the lie. From their existence, we are able to conclude the existence of a lie or lack thereof. This issue needs a bit more attention.

In short, we do not detect lies, as such, at least not detect them directly. We detect the accompanying phenomena (e.g. physiological correlates of emotions) that we observe directly. From their existence or non-existence we are able to make conclusion about emotions, and from the presence of the latter, in the context of the examination (contents of a pre-test interview, contents

²⁰ L. Coleman, P. Kay, *Prototype semantics: The English word lie*, "Language" 1981, 57, pp. 26–44.

²¹ E.E. Sweetser, *The definition of lie: An examination of the folk models underlying a semantic prototype*, [in:] O. Holland (ed.), *Cultural models in language and thought*, pp. 43–66, Cambridge University Press, New York 1987.

²² C.H. White, J.K. Burgoon, *Adaptation and communicative design: Patterns of interaction in truthful and deceptive conversations*, "Human Communication Research" 2001, 27, pp. 9–37.

²³ K. Cantarero, op. cit.

²⁴ R.W. Mitchell, *A framework for discussing deception*, [in:] R.W. Mitchell, N.S. Mogdil (ed.), *Deception: Perspectives on human and nonhuman deceit*, pp. 3–40, State University of New York Press, Albany 1986.

²⁵ A. Vrij, op. cit.

of the test questions), we are able to conclude the existence of a lie or lack thereof. This is the way of lie detection in the case of polygraph testing or noninstrumental techniques consisting in observing behavioural changes (primarily expressional movements: facial expression, pantomime). Another thing, however, is the object of detection techniques based on an analysis of the content of a statement.

Polygraph testing techniques and the object of their detection

Polygraph testing has been routinely used since the 20's of the 20th century as an instrumental technique for detecting lies. It consists in triggering, through a properly conducted pre-test interview, appropriately worded questions of a defined structure, arranged additionally in accordance to a certain scheme, and then observing and registering physiological correlates of emotion with apparatus (polygraph).

Classic technique

The classical technique, also called the Keeler technique or the "relevant-irrelevant" question technique, is the oldest of the routinely used techniques in polygraph testing. It was developed in the 30's of the 20th century by Leonard Keeler²⁶. This technique uses the familiar tests of two types of questions: related to the case (*relevant questions*) and unrelated to the case (*irrelevant questions*). These first questions are sometimes called critical questions, the latter – neutral.

Questions, like all in polygraph testing, must be formulated so that the answers to them could be given briefly: "yes" or "no". So logically they must be "resolved questions"²⁷.

Related questions (critical, *relevant questions*) are supposed to concern the events of the subject matter of the proceedings and eventually the role of the examinee in them. For example, in the case of the murder of a woman, typical questions are the following: "Is this the woman you killed?", "Do you know who killed this woman?", "Were you, on the day of the murder, in this place?" Etc. It is assumed that the examinee, who does not admit to participation in the murder, will answer "no" to these questions. Questions unrelated (neutral, *irrelevant questions*) should concern situations unrelated to the incident which is the subject of inquiry. Such questions concerning the murder of women will be the following: "Is today Wednesday?" (If indeed it is Tuesday), "Are we in Krakow?" (If the test is actually being carried out in Cracow), "Is your name Joe" (if the examinee is actually named Joe and not trying to hide his identity). These questions are formulated so that they themselves have no emotional meaning and that the examinee could answer: "yes".

This technique assumes that the examinee, answering "yes" to questions unrelated, will speak the truth, and answer "no" to questions related (critical), if he was the perpetrator or took any other part in the event under examination, will indicate he lied. The assessment of whether answering "no" to a question (critical), and thus the examinee lied, is the task of the examiner.

Classical technique allowed for next classic tests additionally to use the so-called "peak of tension test" (peak of tension – POT). The essence of such a test is that the critical question is hidden among other apparently similar questions²⁸. For example, in a case where the murder was done with a knife, and the examinee claims not only not to have done this act, but in addition, does not know or has no idea, what tool was used in the murder, may be asked the following questions:

1. Did you hit the woman with a crowbar? (Neutral).
2. Did you kill the woman with an ax? (Neutral).

²⁶ J. Widacki, *Introduction to problems of polygraph testing*, Ed. MSW, Warsaw 1981, p. 56.

²⁷ cf. *Little encyclopedia of logic*, Ossolineum, Wrocław–Warsaw–Cracow 1970, p. 233.

²⁸ J. Widacki, op. cit., p. 62; J.E. Reid, F.E. Inbau, *Truth and Deception. The polygraph (lie-detector) technique*, Williams & Wilkins Comp., Baltimore 1977, p. 55.

3. Did you hit the woman with a vase? (Neutral).
4. Did you stab the woman with a knife? (Critical).
5. Did you shoot the woman? (Neutral).
6. Did you strangle the woman with a rope? (Neutral).

The assumption is that the perpetrator of the act, and in principle only he, will know which question is critical, and answering this one question with the word “no” is a lie. The purpose of the investigator is therefore similar to the conventional tests, to determine if when answering “no” to the critical question, the examinee lied.

Control questions technique

The creator of the control questions technique, based on tests of control questions, is John Reid²⁹. Reid introduced to polygraph testing, in addition to concerning related (critical) and unrelated (neutral) questions, a third type of question: control questions (*control questions*). These are questions thematically regarding the type of the offence under investigation but so worded that the examinee would answer “no”, but in answering, either deliberately lied, or at least was not sure whether or not he lied. The classical layout of the questions in the control questions test by Reid are as follows³⁰:

1. Do they call you, Rudy? [During the pre-test interview the examinee respondent that they usually call him “Rudy”];
2. Are you over 21 years of age? [When the examinee is actually aged 21 years, does not hide it or conceal].
3. Did you shoot John Jons last Saturday?
4. Are we now in Chicago? [If the test takes place in Chicago].
5. Did you kill John Jons?
6. In addition to what you said, did you steal something else?
7. Did you walk to school?
8. Did you steal John Jons watch last Saturday night?
9. Do you know who shot John Jons last Saturday?
10. Have you ever stolen anything from your workplace?

In this test, questions 1, 2, 4, 7 are neutral questions.

Questions 3, 5, 8, 9 are the critical questions, questions 6 and 10 are the control questions.

The control questions examination technique does not compare reactions of critical questions (critical questions) to reactions to irrelevant questions isn't being compared (related) with reactions to irrelevant questions (irrelevant), as takes place in case of those based on the classic tests, but reactions to critical questions are being compared with reactions to control questions. The responses to questions 3 and 5 are compared with the responses to question 6, responses to questions 8 and 9 to the responses to question 10.

The Reid technique involved the construction of two identical control questions tests arranged according to the above scheme, moreover the “mixed questions” test containing all the same questions as the control questions test, but posed in another previously unknown order. Depending on needs and possibilities, as part of this examination technique, POT tests as well as supporting “yes” (Yes-test) test or the control questions test in the “silent version” (Silent Answer Test - SAT) may be performed as part of this technique. In the first of these tests, we asked the examinee that regardless of whether he would answer true or false, for every question answer “yes”. According to the recommendations of

²⁹ J.E. Reid, *A revised questioning technique in lie-detection test*, “Journal of Criminal Law, Criminology and Police Science”, 1947, 37, pp. 542–547.

³⁰ Ibidem, p. 31.

Reid-Inbau³¹ this test is performed when the basic control questions test questions, including questions of the mixed test, gave no grounds for a determination but gave suspicion that the examinee during the study tried to manipulate their reactions. It is assumed that such a person is in the test who agree, will attempt to artificially induce a clear response to the critical questions in order to confuse the examiner. In the “silent version” the examinee is instructed not to answer aloud, only to answer in his mind, but truthfully. The “silent response” test, as the result of the practice, gave better results with certain personality types than the ordinary control questions test, during which the examinee must say aloud, “yes” or “no”.

With the control questions technique, Reid, as an auxiliary diagnostic criteria, in addition to the assessment of the response recorded by the polygraph, incidentally took into account the assessment of the potential interference with the test records, verbal and nonverbal behavioral symptoms associated with the pre-test, during and after the test³².

On the basis of all the conditions, responses to the critical questions (in relation to the response to the control questions), responses from the POT test, ancillary tests, assessing potential storage disorders, behavioral assessments (through the prism of verbal and non-verbal symptoms), the examinee belongs to one of following three groups: deceptive (*deceptive*), non-deceptive (*non-deceptive*), or inconclusive (*inconclusive*).

In the correct methodologically opinion, the examination should be written: **“The examinee responds to critical test questions as usual individuals who answer such questions deceptively, that is, they are lying or concealing the fact of having information related to the case”**.

While the answers to individual critical questions of the control questions test (along with the context of the question) may be statements logically false and, therefore, speak to the examiner of the intention that the statements recognized to be true, are lies, then **in the responses to each question were not assessed separately. Evaluation of the sincerity of the examinee is made as a whole, on the basis of the responses to all questions. It does not detect individual lies to individual critical questions, but generalizes the assessment of the whole test, classifying the test as “deceptive” or “non-deceptive” (or, in the absence of a clear reaction, the result is considered inconclusive)**. Not evaluation, e.g. “the examinee lied, answering question 5”, “examinee did not lie, answering question 9”, but all reactions from the test are averaged, ranking the examinee to one of the above mentioned groups: “non-deceptive” (*non-deceptive* – NDI), “deceptive” (*deceptive* – DI), or “inconclusive” (*inconclusive* – INC). In addition, in the “silent testing”, which incidentally are also used in the evaluation, in general, we are not dealing with a lie in the above sense, because the examinee does not say anything knowingly and intentionally does not show.

Identical in this respect is in the control questions technique arising from the development of the techniques of Reid, and thus the Backster Zone Comparison Test or newer, and also for the Federal Zone Comparison Test, Utah Zone Comparison Test³³.

Thus, polygraph testing performed using the control questions technique does not detect a lie in response to particular questions of the test, but after all responses to individual critical questions, compared with reactions to control questions, are brought to a whole, whether the test, during the test is non-deceptive or deceptive, or more precisely, attempted to mislead the examiner, or not³⁴, This applies both to qualitative methods of the analysis of records, as well as quantitative (numerical)³⁵.

³¹ Ibidem, pp. 32–33.

³² J. Widacki, *The analysis of diagnostic premises in polygraph examinations*, Ed. University of Silesia, Katowice 1982, pp. 78–84.

³³ J. Widacki (eds.), *Criminalistic*, C.H. Beck, ed. 2, Warsaw 2012, p. 383.

³⁴ cf. J. Widacki (Eds.), *Polygraph testing in Poland*, Publishing House AFM, Cracow 2014, p. 196.

³⁵ cf. for example. J.A. Matte, *Forensic psychophysiology. Using the polygraph*, J.A.M. Publications, Williamsville, New York 1996, p. 398 et seq.; D.C. Raskin, Ch. Honts, J.C. Kircher (ed.), *Credibility assessment. Scientific research and applications*, Elsevier – Academic Press, 2012, pp. 4–7, 161–210, 305–307.

It is worth paying attention to the professional scientific literature, from the United States at least since the 50s of the 20th century, they do not speak about a lie (*lie*) as the object of polygraph testing, but about “misleading”, “insincerity”, “deception” (*deception*)³⁶. The word “*deception*” in the English language it is understood as a deliberate act to cause someone to believe in something that is not true³⁷, or the use of fraud, deception, misrepresentation³⁸, or a deliberate act or omission to provide information with the intention of misleading³⁹.

Lykken technique

In the years 1958/1959 Lykken conducted two experiments that aimed, he said, not to detect “lies”, but “guilty knowledge” (*guilty knowledge*)⁴⁰. In the Lykken experiments, questions were not asked according to the techniques of the control questions test technique or the classical technique. In every test, and a few of them were conducted, they were setting one, preceded by the specific introduction, entering the question, reading alternative replies later, from which one was really associated with the directed criminal event. The first of these questions was, “If you are the murderer, you will know that there was an unusual object present in the murder room. One of the following:

1. tape recorder?
2. easel?
3. box of chocolates?
4. chess piece?

The examinee, to these questions did not answer, but listened, and a psychogalvanometer recorded their reactions. To strengthen the motivation of the respondents, additional electric shock electrodes, indicating that when the GSR change pointed to guilty, they would receive a painful shock⁴¹. As you can see in the Lykken experiments, the examinees did not respond to questions, so it is difficult to speak about “lie” (*lie*) and “deception” (*deception*).

The object of detection in polygraph testing performed using the control questions technique and the Lykken technique

It has been said that the control questions technique aimed to detect lies and the Lykken technique (Guilty Knowledge Test – TCH, Concealed Information Test – CIT) aimed at detecting “guilty consciousness”, “hidden information” or “knowledge of the deed”. Indeed, are these techniques used for a detection of something else? Is there another object of recognition?

The literature points out **that in both the control questions technique and the Guilty Knowledge Test technique, the object of detection is essentially the same thing⁴². It is the – variously called –**

³⁶ cf. e.g. C.D. Lee, *The instrumental detection of deception*, Ch. Thomas, Springfield, 1953; J.E. Reid, F.E. Inbau, op.cit.; Ch.R. Honts, *Psychophysiological detection of deception*, Current Directions in Psychological Sciences 1994, pp. 77–82; J. Vendemia, *Detection of deception*, “Polygraph” 2003, 32, pp. 97–106; D.C. Raskin, Ch.R. Honts, J.C. Kircher (ed.), op. cit., pp. 86–88.

³⁷ *Oxford advanced learner's dictionary*, Oxford Univ. Press 2010, p. 391.

³⁸ *The American Heritage dictionary*, Houghton Mifflin Comp., Boston 1985, p. 371.

³⁹ D. Krapohl, S. Sturm, *Terminology reference for the science of psychophysiological detection of Deception*, APA, Chattanooga. 27–29, 1997, p. 20.

⁴⁰ D.T. Lykken, *The GSR...*, op. cit.; D.T. Lykken, *The validity...*, op. cit.

⁴¹ D.T. Lykken, *The GSR...*, op. cit. A more late technique of polygraph testing called the Guilty Knowledge Test technique (GKT) or Concealed Information Test (CIT) is based on these experiments.

⁴² J. Widacki, *On the choice of Polygraph testing technique. Is technique based on GKT tests (CIT) is better than techniques based on CQ tests?* “Problems of Criminalistics” 2011, 273, pp. 5–10; J. Widacki, *Logical identity of Conclusions from polygraph testing performer in Control Question Test (CQT) and Guilty Knowledge Test (GKT) techniques*, “European Polygraph”, 2011, 1 (15), pp. 1–10.

insincerity of the examinee, diagnosed after the entire examination, and not the insincerity of the responses to individual test questions.

What is the difference between saying, “The examinee’s answer to critical questions is insincere”, (as we say in the Control Questions Test – CQT) and saying, “The examinee is concealing his knowledge of the facts of the case”, (as we say in the Guilty Knowledge Test – GKT)?

In polygraph testing carried out for the purposes of a criminal trial, the principle of the sincerity/insincerity of the reply is subject to an evaluation of the examinee concerning a determined thematic area and limited test questions, rather than the sincerity/insincerity of the reply to each individual question. In contrast, in pre-employment examinations (*pre-employment*) an assessment of the grey answers to particular questions may be significant.

Other instrumental methods of lie detection

In addition to polygraph testing, which uses the polygraph to register and enable the assessment of physiological correlates of emotions such as changes in respiration, changes in heart rate, changes in the galvanic skin reflex, as well as changes in plethysmographic response, there are attempts to use other apparatuses for lie detection, recording other physiological correlates of emotions or allowing the observation of processes in the brain.

Attempts have been made to detect lies by observing changes in voice, using a special device, such as the “Psychological Stress Evaluator” (PES), “Voice Stress Analyzer” (VSA) and “Computerized Voice Stress Analyzer” (CVSA)⁴³. There are also attempts of lie detection on the basis of facial temperature changes recorded by the thermal imager⁴⁴ or evaluation of eye movements recorded by a specially designed device for this purpose. All these methods aimed or are heading at identical targets and purposes as polygraph testing, that is the detection, as identically as in polygraph testing, of uncovering deception, concluded from the evaluation of physiological changes of stimulated correlates of emotion with appropriately worded questions.

Instrumental lie detection looks differently and is based not on the induction and observation of physiological correlates of emotion, but on the observation of brain processes. The first of these is electroencephalography. The second is the observation of brain magnetic resonance imaging (fMRI). The first experimental attempts to detect lies using EEG were taken by Obermann in the late 30s⁴⁵. And then continued in the 70s of the 20th century, among others, by Dufka and Richter⁴⁶, Guljajewa and Bychowski⁴⁷. In recent years, beginning in the late 90s of the 20th century, there was a renewed interest in the use of EEG to detect lies⁴⁸.

In recent years, many experimental studies were made involving the use of magnetic resonance imaging (fMRI) to detect lies⁴⁹. Without going into it at this point in the discussion on how these tests can (and possibly when) replace traditional polygraph testing, it is clear that they are generally identical to lie detection tests in polygraph testing (conscious expression contrary to the truth and to

⁴³ J. Widacki (ed.), *Forensics*, op. cit., p. 393.

⁴⁴ cf. I. Pavlidis, N.L. Eberhardt, J.A. Levine, *Seeing Through the Face of Deception*, “Nature” in 2002, 415, pp. 35–35; D.Q.A. Pollina, F. Horvath, J.W. Denver, A.B. Dollins, T.E. Brown, *Development Technologies and Test Formats for Credibility Assessment*, “European Polygraph” 2009, 3, pp. 3–4.

⁴⁵ C.E. Obermann, *The effect on the Berger rhythm of mild affective states*, “Journal of Abnormal and Social Psychology”, 1939, 34, pp. 84–95.

⁴⁶ M. Dufek, L. Richter, *Soudne psychologicka praktikum*, Universita Karlova, Praha 1972, p. 20; M. Dufek, *K issues polygrafického vysetrovani v kriminalistice*, in:] *Doplňkové study materials pro kriminalistický smer právnického studies*, Universita Karlova, Praha 1970, pp. 24–25, 55–57, 88–92.

⁴⁷ P.J. Guljajew, J.E. Bychowski, *Issledovanie emotsionalnogo sostojanija cziefowieka in Processia proizvodstva sledstviennowo diejstwija*, “Kriministika and Sudiebnaja Ekspertiza” 1972, 9, p. 108.

⁴⁸ cf. J. Wojciechowski, *Detection of concealed with information of the P-300 potential amplitude analysis*, “European Polygraph” 2014, 8, 4 (30), pp. 167–188.

⁴⁹ cf. J. Vendemia, *fMRI as a method of detection of deception*, “European Polygraph” 2014, 8, 1 (27), pp. 5–21.

mislead). **The fundamental difference between EEG examinations or fMRI and traditional polygraph examinations consists – except that it uses a different apparatus – in that these studies do not involve drawing conclusions about deception from the results of the analysis of physiological correlates of emotions that are secondary processes in relation to changes occurring in the brain, but the observation and analysis of the latter.**

Lie detection based on the assessment of behaviour and linguistic aspects of speech

For ages, psychologists have paid attention to the fact that the inference of deception based on so-called non-verbal signs of behaviour is burdened with many mistakes⁵⁰. In the case of the observation of behaviour, it isn't possible to talk about the so-called Pinocchio's nose, a universal behaviour for all people which clearly indicates a lie or insincerity. The only things that can be seen, for example, are indicators of specific emotional states, which relate to deception. So we can say that in the case of behavioral clues of deceit, the situation is similar to polygraph testing: evaluation is subject to correlates of emotion. In the case of the polygraph – physiological, observed (and recorded) through the apparatus, and thus “instrumental”. In the case of methods based on an evaluation of behavior, it is observed directly without the use of the apparatus.

Proponents of behavioral methods indicate that behaviour is a trait that the communication of, is not subject to large voluntary control. Persons seeking to mislead, focus more on what they say (to tell the lie convincingly and consistently) than on how to say or how to behave at the same time. The main hypothesis underlying the evaluation of behavior in order to seek insincerity through expression is the hypothesis of Zuckerman, DePaulo and Rosenthal⁵¹. According to it, people who lie, are in a different situation than those who speak the truth. Firstly, they experience high cognitive load because they need to formulate a probable lie, enriching the lie in detail, making it plausible and even memorizing it so that their account is consistent over time. Second, they experience other emotions than people who speak the truth: the guilt associated with lying to someone, fear that a lie will be revealed through the excitement associated with the deception. Thirdly and finally, liars try very much to control their behaviour in order to make a positive impression on the person being deceived. On the basis of this hypothesis, it is concluded that the above characteristics of the situation in which the liar stands, translate into behavior, i.e. the act of lying usually takes some intellectual effort, feelings of a sense of guilt or anxiety during lying and be indicated through various behaviours. In other words, in the case of behavioural indicators, the direct detection of lies is not taken into account, only the detection of indicators showing the cognitive load, experiencing certain emotions or control of behaviour. As claimed by some researchers, in the case of behavioral analysis judging whether the single message is sincere or not⁵², draws attention to the fact that the analysis depends on the whole testimony or statement⁵³.

To date, there are several dozen individual indicators of lies (e.g. visual: looking in the eyes, the presence of illustrating gestures whether vocal: level of the voice, pauses in speech). Unfortunately, relying only on these guidelines does not lead to accurate estimates⁵⁴. Researchers point out that perhaps the behavioural pointers alone are not sufficient to accurately evaluate the sincerity of a statement.

⁵⁰ e.g. B.M. DePaulo, J.J. Lindsay, B.E. Malone, L. Muhlenbruck, K. Charlton, H. Cooper, *Cues to deception*, “Psychological Bulletin” 2003, 129 (1), pp. 74–118; S.L. Sporer, B. Schwandt, *Para verbal indicators of deception: A meta-analytic synthesis*, “Applied Cognitive Psychology” 2006, 20, pp. 421–446; V. Hauch, I. Blandon Gitlin, J. Masip, S. Sporer, *Are computers effective lie detectors? A meta-analysis of linguistic cues to deception*, “Personality and Social Psychology Review”, 2014, pp. 1–36.

⁵¹ M. Zuckerman, B.M. DePaulo, R. Rosenthal, *Verbal and nonverbal communication of deception*, [in:] L. Berkowitz (ed.), *Advances in experimental social psychology*, 14, pp. 1–59, Academic Press, New York 1981.

⁵² e.g. *On the basis of microexpression*, P. Ekman, Facial expressions, [in:] C. Blakemore, S. Jennett, (ed.), Oxford Companion to the Body, Oxford University Press, London 2001.

⁵³ A. Vrij, *Detecting...*, op. cit.

⁵⁴ cf. B.M. DePaulo, J.J. Lindsay, B.E. Malone, L. Muhlenbruck, K. Charlton, H. Cooper, *Cues...*, op. cit.

According to the Interpersonal Deception Theory⁵⁵, deceit or deceptive testimony is an act of communication and personal behaviour of the recipient of the deception (interviewers) has a very large impact on how it will be presented and how easy it will then be detected. Because deception (disingenuous testimony) is often a dynamic process of creating the appropriate context (i.e. involvement in interaction, strategic actions, creating a relevant conversation, etc.) by the recipient to facilitate fraud detection. One of the examples of the creation of a systematic evaluation of behaviour on the basis of the behavioural indicators linked to the appropriate way of questioning is ACID⁵⁶ or the Behavioural Analysis Interview (BAI)⁵⁷, however, its effectiveness is also debatable⁵⁸.

An altogether different issue is assessing the sincerity of behavior on the basis of linguistic analysis⁵⁹. The basis of linguistic indicators of deception is the aforementioned concept of Zuckerman, DePaulo, and Rosenthal. It also draws attention to the fact that a lie is created differently than a true account. When a person testifies honestly, he uses his long-term memory and most likely also his autobiographical memory. The liar would rather benefit from semantic memory and specific cognitive processes, rather than a person that speaks the truth⁶⁰. All of this will translate into another way of expressing the insincere statements of liars, e.g. different choice of words, used in a different way, especially all embracing, another level of refinement of expression etc. Until now, more than 200 linguistic clues have been identified that could directly indicate insincere testimony⁶¹. It is worth noting that linguistic analysis takes into account all of the testimony, not excerpts. In other words, in the case of linguistic analysis one cannot talk about the detection of lies, defined as a single statement, but rather of assessing the sincerity of the entire expression (the entire testimony) on the basis of an analysis of its verbal characteristics.

Lie detection on the basis of an analysis of the content of the statement

One of the directions of research into the detection of lies/insincerity in testimonies is related to trying to answering the question of whether it is possible to find, in the content of the testimony of the witnesses, the deliberate falsification of elements and what that could be. In the beginning, they tried to set different criteria and methods for the assessment of the content of the testimony (apart from that of the witness) that would indicate the existence of insincere, spurious⁶² elements. And later they sought to group these criteria in a uniform method based on a systematic analysis of the evidence in order to find the so-called insincerity criteria⁶³.

In the case of methods based on the analysis of the content of speech, it is difficult to talk about the direct detection of lies because these methods are often based on the so called Undeutsch hypothesis⁶⁴ whereby,

⁵⁵ D.B. Buller, J.K. Burgoon, *Interpersonal Deception Theory*, "Communication Theory" 1996, 6, pp. 203–242.

⁵⁶ K. Colwell, C.K. Hiscock-Anisman, A. Memon, L. Taylor, J. Prewett, *Assessment Criteria Indicative of Deception (ACID): An integrated system of investigative interviewing and detecting deception*, "Journal of Investigative Psychology and Offender Profiling" 2007, 4, pp. 167–180.

⁵⁷ F.E. Inbau, J.E. Reid, J.P. Buckley, B.C. Jayne, *Criminal interrogation and confessions*, Gaithersburg, MD: Aspen 2001.

⁵⁸ A. Vrij, S. Mann, R.P. Fisher, *An empirical test of the Behaviour Analysis Interview*, "Law and Human Behavior" 2006, 30, pp. 329–345.

⁵⁹ Or verbal clues lie, cf. A. Vrij, *Detecting...*, op. cit.

⁶⁰ I. Blandon-Gitlin, E. Fenn, J. Masip, J., A. Yoo, *Cognitive load approaches to detect deception: Searching for cognitive mechanisms*, "Trends in Cognitive Sciences" 2014, 18, pp. 441–444.

⁶¹ V. Hauch, I. Blandon-Gitlin, J. Masip, S. Sporer, *Are computers...*, op. cit.

⁶² e.g. F. Arntzen, *Psychology testimony of witnesses*, PWN, Warsaw 1989.

⁶³ M. Steller, G. Kohnken, *Statement analysis: Credibility assessment of children's testimonies in sexual abuse cases*, [in:] D.C. Raskin (ed.), *Methods in Psychological criminal investigation and evidence*, pp. 217–245, Springer, New York 1989.

⁶⁴ U. Undeutsch, *Beurteilung von der Glaubwürdigkeit Zeugenaussagen*, [in:] U. Undeutsch (ed.), *Handbuch der Psychologie*, Band 11: Forensische Psychologie, Göttingen: Hogrefe 1967.

- a. statements based on personal experience of the witness will differ in terms of the form and contents (qualitatively and quantitatively) from insincere statements and
- b. the person testifying honestly will differ from those who are insincere in terms of motivation, which translates into the content of the testimony.

It also indicates that according to the model of reality monitoring (Johnson, Raye, 1981), memories of real events differ from imagination and false memory. True memories of real personal experiences reflect perceptual processes occurring in the course of their acquisition (e.g. contain information about smells, tastes, colours etc.), and emotions and memories based on ideas contain more reasoning and cognitive processes. This model has been adapted by researchers for the assessment of the evidence on the need to create tools to detect insincere testimony (Reality Monitoring – RM, Reality Monitoring)⁶⁵.

It is easy to figure out methods that are created on the basis of the aforementioned hypotheses can indeed differentiate between testimony based on an actually survived factual event and the testimony of the completely imaginary, but do not allow the detection of which **items** are inaccurate events (inconsistent with reality) or deceptive (as to which elements of the testimony are made up). It is not difficult to imagine a situation in which an honest witness tells a true account, e.g. bank robbery, but conceals some information (e.g. small elements of his behavior) or inaccurately recalls some things (e.g. appearance of the perpetrators). In this case, the use of methods based on an analysis of speech content can be problematic, because it will only indicate that the viewer actually survived a bank robbery but it will be very difficult to point out which parts of his speech are intentionally falsified. In other words, as similar to the case of linguistic methods or other methods based on the analysis of behaviour, in methods based on analysis of a statement, the detection of an isolated lie is not an object of interest as to a single detail or – as in a polygraph examination – to an area limited by the test questions, only the evaluation of the frankness of the witness, understood as related to the event which indeed was survived.

Therefore, initially such methods as CBCA (Criteria Based Content Analysis, based on the criteria of evaluation of evidence) were used only to evaluate the testimony of sexually abused children⁶⁶, and now it can be said that these methods are useful even in the testimony of children – victims of other types of crime⁶⁷, children witnesses⁶⁸, adult witnesses⁶⁹, adult victims⁷⁰ and offenders⁷¹.

To sum up, in the case of methods based on the analysis of the content of testimony, it is difficult to talk about “lie detection” (as in the case of polygraph testing); the term “assessment of the sincerity of testimony” seems more accurate or, more precisely, the “assessment of how likely it is that the witness survived the event in the broad outline described”.

One of the most frequently used methods of assessing the sincerity of the content of the testimony by experts in the US, Germany and the Netherlands, is the SVA (Statement Validity Analysis, Quality

⁶⁵ S.L. Sporer, *Reality monitoring and detection of deception*, [in:] P.A. Granhag, L.A. Strömwall (ed.), *Deception of detection in forensic contexts*, Cambridge University Press, Cambridge, 2004, pp. 64–102.

⁶⁶ Mainly in order to determine whether abuse has actually taken place – e.g. A. Craig, R. Scheibe, D.C. Raskin, J.C. Kircher, D. Dodd, *Interviewer questions and content analysis of children's statements of sexual abuse*, “Applied Developmental Science” 1999, 3, pp. 77–85.

⁶⁷ K. Pezdek, A. Morrow, I. Blandon-Gitlin, G.S. Goodman, J.A. Quas, K.J. Saywitz, *Detecting deception in children: Event familiarity affects Criterion-Based Content Analysis ratings*, “Journal of Applied Psychology” 2004, 89, pp. 119–126.

⁶⁸ A. Vrij, L. Akehurst, S. Soukara, R. Bull, *Let me inform you how to tell a convincing story: CBCA and Reality Monitoring Scores as a function of age, coaching and deception*, “Canadian Journal of Behavioural Science”, 2004, 36, pp. 113–126.

⁶⁹ Ibidem.

⁷⁰ K. Landry, J. Brigham, *The effect of training in Criteria Based Content Analysis on the ability that detect deception in adults*, “Law and Human Behavior” 1992, No. 16, pp. 663–675.

⁷¹ S. Porter, J.C. Yuille, D.R. Lehman, *The nature of real, implanted, and fabricated memories for emotional childhood events: Implications for the recovered memory debate*, “Law and Human Behavior” 1999, 2, pp. 517–537.

Evaluation Testimony) procedure⁷². This procedure consists of four parts. Firstly, the expert acquaints himself with the case file, mainly to look for information about the witness and the event and to pre-analyse the testimony. The second part is a structured examination of the witness, which should be recorded and then transcribed because the next part of the procedure relates solely to the analysis of the content of speech. Analysis carried on video, for example, is wrong because the assessment of an expert can affect the behavior of the witness, his appearance or other non-substantive factors. Also incorrect is to conduct analyses the content of testimony based on the record of hearings, because in general they do not contain accurate witness statements, but only reformulated and drawn up by the examining person. Moreover, the scope of the statement of the examinee, strengthened in the form of a protocol, is often stimulated with questions of the examiner which are not recorded in the protocol.

The third element of the SVA procedure is Criteria Based Content Analysis (CBCA previously mentioned), and the last, the so-called Validity Checklist, which works in testimony obtained from psychological characteristics (e.g. if the witness speaks appropriate to their age and knowledge), the characteristics of the hearing itself (e.g. if it was not conducted in a prompted manner), aspects of incentive (e.g. why the witness decided to give evidence), and compatibility with other testimony or laws of nature.

In the words of the developers⁷³, the most important step in the SVA is the interrogation, and then the CBCA. The CBCA itself consists of 19 criteria related to the semantic features of the content of the testimony, which more often appear in non-deceptive testimony (concerning the events of which were actually witnessed) than deceptive (on imaginary events). These criteria are, for example, the logical structure of the statements (criterion 1), number of details (criterion 3) or reconstruction of the statements (criterion 5). It is assumed that the testimony of the true criteria will be of a greater intensity than a false testimony. More specifically, in the testimony of sincere people, those trained in the use of CBCA methods should more often state the existence of the logical structure of the statements, diagnose a greater number of details or more likely to state the existence of a statements or parts of conversations cited in their original form than in imaginary testimony. The 19 truth criteria were identified on the basis of many years of experience and observation and have been verified in numerous scientific studies. In general we can say that most of the criteria actually occur more frequently and with greater intensity in sincere testimony, i.e. on experienced, true events⁷⁴.

A similar, though not such comprehensive method for assessing the content of the testimony, is the Reality Monitoring (RM). This tool is comprised of 8 or 7 criteria⁷⁵. These criteria are based on well-described and validated psychological theories concerning the formation of memories and relate, for example, spatial information (in a sincere testimony there should be more), transparency (sincere testimony should be more lively and transparent) and feelings (the more the emotions described in testimony, the more likely it is based on the actual survived event). Although it may seem that methods such as SVA or RM, are fairly easy to learn, the researchers suggest that in order to be proficient in seeking out the criteria requires lengthy training⁷⁶.

Evaluating testimony using CBCA or RM is more or less the same way and is usually done in two phases. In the first phase, the expert seeks criteria for the truthfulness of the testimony, and then assigns a point value. It is usually a 0 – if the criterion is not present, 1 – if there is a low intensity or 2 – if strongly present. Then, as is done in experiment studies, the points of the testimony are summed up and using statistical analysis compares with results obtained from sincere and insincere testimony, determining whether there are significant differences between them. It should be emphasized, that in

⁷² M. Steller, G. Kohnken, *Statement...*, op. cit.

⁷³ G. Kohnken, *Statement Validity Analysis and the detection of the truth*, [in:] P. A. Granhag, L.A. Strömwall (ed.), *The detection of deception in forensic contexts*, Cambridge University Press, Cambridge, 2004, pp. 41–63.

⁷⁴ R. Volbert, M. Steller, *Is this truthful testimony, fabricated, or based on false memory? Credibility assessment 25 years after Steller and Köhnker*: (1989), "European Psychologist" 2014, pp. 207–220.

⁷⁵ S.L. Sporer, *Reality...*, op. cit.

⁷⁶ V. Hauch, S.L. Sporer, S.W. Michael, Ch. A. Meissner, *Does training improve the detection of deception? A meta-analysis*, "Communication Research" 2014.

practice, this method does not apply – simply, the expert has at his disposal only one testimony, on which it must be considered whether it is genuine or not. Since, for none of the above described analysis methods of content there is no objectively placed border of at how many points sincere testimony starts, the expert based on the acquaintance of the tools, correctly performed analysis, psychological knowledge and other factors must alone make the decision as to the evaluation of the truthfulness of a given testimony.

It is worth noting that a problem also occurs in instrumental methods (such as polygraph or thermography) and other non instrumental methods. For each of these techniques the person making the analysis of sincerity/truth of the account on the basis of indication methods (the truthfulness criteria for the CBCA, verbal behaviour sincerity criteria in the case of behaviour analyses, data from three channels polygraph etc.) must assess whether the statement is true or not. This means **that even if the indicators are well researched and fairly well demonstrate the genuineness of the account, its conclusions, to the persons using the data methods, could be wrong. The reason may be most likely the fact that professionals using the methods of determining the sincerity of testimony or detecting lies, making a final decision as to whether a statement is true or false do not always follow the indications of the method, only their own convictions and nonsubstantive premises.**

Rating the so-called psychological credibility

Psychological credibility is not a true method of detecting or estimating the truthfulness of testimony; however, it is an issue which is directly associated with such methods. In practice, it quite often happens that expert psychologists speak of the issue of “credibility” and “truthfulness” of the witness or sometimes even the accused. It happens that they are even encouraged to do so by the judiciary. Apart from the formal consideration that the credibility (or truthfulness) of the witness testimony is reserved for the judiciary and in this respect the expert cannot replace the judiciary, it must be stated clearly that the psychologist does not have the tools to make such an assessment. **So whenever an expert psychologist talks about the credibility of the witness, he infringes upon criminal procedure, speaks contrary to the knowledge of psychological but on their current knowledge and life experience, and runs beyond the border of “special message” of which is referred to in Art. 193 of the Polish Code of Criminal Procedure**⁷⁷. An expert psychologist can only assess, a priori, the credibility of the witness, taking into account such elements as personality traits (e.g. susceptibility to suggestion or a tendency to manipulate people), the ability to perceive and recount observations, a tendency to confabulation (or fill in memory gaps with bogus content) etc.

All of these features can facilitate the testimony of the subjectively real, but objectively false. Thus, their statement may be an important clue to facilitate the judicial body’s assessment of such evidence, in particular, suggesting caution in that assessment. Failing to state such characteristics is sometimes called “psychological credibility”⁷⁸. This “psychological credibility” is obviously something completely different from the reliability of the specific statement of a particular witness. Plainly speaking, psychological credibility of the witness (as sources of information) is limited to the conclusion that the given witness, if he does not want, isn’t lying. **That is, a “credible psychologically” witness is a witness in whom there is no impairment of perception, memory, reconstruction of observations or such personality traits that could support giving evidence objectively false, though subjectively true.**

But in practice you can meet different understandings of “psychological credibility” that, the expert in an unauthorized manner, does not refer to the source, the witness, but his testimony. So for example, in one case, an expert psychologist wrote, “From a psychological point of view, the content of the testimony of the minor JG show signs of credibility”⁷⁹. As is clear from the supplementary oral

⁷⁷ cf. J. Widacki, *Attempts to verify the truthfulness of the criminal process*, “Palestra” 2012, 3–4, pp. 13–18.

⁷⁸ Ibidem, p. 14.

⁷⁹ Case 2 from 1386 / 13 of Regional prosecutor’s office in K.N.H.

opinion, the basis for such a claim of the expert was to establish in the psychological examination, “no basis for finding in the examinee a tendency of confabulation or memory disorders”.

As once again it is worth noting that the ability of a witness with appropriate predispositions to testify (properly developed personality or ability to reconstruct observations) does not determine that his testimony will be sincere. A witness may indeed have the intention of misleading the judicial body and testifying insincerely. Therefore, expert psychologists sometimes use the term **“hallmarks of psychological credibility”** to determine whether the evidence bears the so-called psychological signs of credibility. They are most often understood or identified with the above mentioned content criteria of frankness (mainly taken from CBCA) or described as resulting from the process of forming testimony (SVA). In other words, the expert, describing the psychological credibility of the testimony, **refers not to the characteristics of the witness, only to the testimony of the specific characteristics of a particular case** – and more specifically whether he can find the criteria of sincerity. We can not forget that the diagnostic value of formalized procedures such as SVA and CBCA is far from 100%. **Therefore, use of a non-standard method, CBCA or SVA, can give only general, informative, indicative, never a conclusive result. It is more the result of the personal conviction of the expert than the result of a measurement.**

However, even the so-called indicative result of the analysis, the so-called psychological credibility of the testimony of a witness, or testimony assessment process evaluation (how and under what circumstances it was made), the variables affecting the formation of testimony, and as a consequence – the possible criteria for indicating the honesty / dishonesty of the testimony, may be helpful to an experienced judge, understanding the limitations of the method and facilitating its assessment of the probative value of the evidence submitted.

Summary

As it should be stressed, in criminology and forensic psychology the words “lie” and “lie detection” function in different senses, the object of “lie detection” is also different. The table below shows a summary of detection methods currently functioning along with a character sketch of assumptions. As shown in the above table, at present, existing tools, determined often as being used for a detection of the lie, in fact don’t directly measure them at all. In most cases, it is assumed that (1) the lie causes changes in emotional, cognitive, etc. and that (2) the change in emotional and cognitive impact on the observable characteristics (physiological change, behavior change etc.), and as a result, on the basis of observable characteristics, inferences can be made on the existence of deception in a statement or an entire testimony. In the event that one of the assumptions is not met (e.g. the person lying does not feel any emotional connection to it), the use of such tools ceases to make sense, hence the great enthusiasm associated with the idea of using fMRI in the study of deception. In principle, proponents of this method directly observe the formation of the lie in the place where it occurs (e.g. specific brain structures)⁸⁰ rather than inferring a lie based on observing its supposed correlates. In addition, the use of fMRI by its followers allows the estimation of whether individual simple statements (“I did not kill him”) or words spoken, are disingenuous. It should be noted, however, that this method at its present state of knowledge is still very unreliable.

It is worth once again emphasizing that currently used practice methods of lie detection in their essence are dealing with observable and established correlates of lies and allow only a cautious inference about the sincerity as a whole of a specifically defined statement, whereby there seems to be no doubt that the diagnostic value of the polygraph test is, of all these methods, the

⁸⁰ It should be noted that this statement is a gross simplification, cf. E. Rusconi, T. Mitchener-Nissen, *Prospects of functional magnetic resonance imaging as lie detector*, “Frontiers in Human Neuroscience” 2013, 7, pp. 1–12; cf. also J. Vendemia, *fMRI as a method of detection of deception: a review of experience*, “European Polygraph” 2014, 8, 1 (27), 5–21.

highest, comparable to the diagnostic value of others routinely used in forensic identification methods, sufficient for any indirect evidence.

Table 1. Summary of lie detection methods and correlates examined by them

	CQT techniques, GKT and techniques based on an assessment of physiological correlates of emotion	Classical technique (Keeler, IR)	Techniques based on an assessment of brain function (e.g. study using fMRI or EEG)	Techniques based on an analysis of nonverbal behavior	Techniques based on content analysis (e.g. CBCA, RM, SCAN) and linguistic indicators
Correlative lies	Emotional changes	Emotional changes		Emotional and cognitive changes	Cognitive and motivational processes
observable traits of the subject	Physiological changes	Physiological changes	Changes in brain function	Change in non-verbal behavior	Content of statements
Result of using methods	The overall assessment of the sincerity of the subject during the test	Evaluation of insincerity in answers to individual critical questions	Assessment of the sincerity of a single sentence	Assesses the sincerity of the statement	An assessment of whether the testimony is based on the actual experienced event

Source: Tab. 1: authors.

Article carried out within the framework of the project DEC-2013/11/B/HS5/03856 funded by National Center for Research and Development

Translation Ronald Scott Henderson

Jan Widacki

Poligraf czy wariograf?*

Nazwa „wariograf” wprowadzona została przez prof. Pawła Horoszowskiego. Po raz pierwszy użył tej nazwy w opinii z badań, przeprowadzonych 27 czerwca 1963 roku, w sprawie o zabójstwo prowadzonej przez Prokuraturę Wojewódzką w Olsztynie (sygn. II Ds. 25/63). Natomiast w literaturze naukowej, po raz pierwszy nazwy „wariograf” Horoszowski użył dopiero w roku 1965 (Horoszowski, 1965). Wcześniej posłużył się nią w popularnonaukowej książce *Od zbrodni do kary* (Horoszowski, 1963).

Tak naprawdę, jako pierwszy nazwy tej użył prokurator w postanowieniu o powołaniu biegłego, zlecając biegłemu „przeprowadzenie badań wariograficznych”. Z całą pewnością, nie on jednak ją wymyślił, a fakt, że użył jej w postanowieniu świadczy o tym, że treść postanowienia konsultował wówczas z powoływanym właśnie biegłym.

Wcześniej, na przykład w wydanym w 1958 roku podręczniku *Kryminalistyka*, Horoszowski posługiwał się nazwą amerykańską „lie-detector”. Jako ciekawostkę można podać, że już dawno próbowano wprowadzić w literaturze polskiej nazwę „odkłamywacz” (Kreuz, 1949), ale nazwa ta nie przyjęła się.

Nazwa „wariograf” używana jest w polskiej terminologii, tak zawodowej, jak i naukowej, zamieniane z nazwą „poligraf”. Ta ostatnia, w rozmaitych wariantach językowych (ang. „polygraph”, niem. „der Polygraf”, ros. „poligraf”) używana jest na całym świecie. Nazwa „wariograf” stosowana jest tylko w Polsce.

Obydwie nazwy są w zasadzie synonimiczne. „Poligraf” jest nazwą grecką (wywiedzioną z *poloi* – wiele, *graphos* – pisać). Z kolei w nazwie „wariograf” pierwszy człon grecki zastąpiono wyprowadzonym z łaciny „vario” (od *varius* – różny), pozostawiając grecką końcówkę. Nawiasem mówiąc identyczną etymologię ma słowo „wariat” („różny”, „różniący się” od normalnych).

Osobiście w używaniu nazwy „wariograf” nie widzę nic złego, podkreślając wielokrotnie, że nie jest ważne, czy badanie nazwiemy „poligraficznym”, czy „wariograficznym”, byle by było ono wykonane zgodnie z regułami sztuki (Widacki, 2014). Dla niektórych, z niezrozumiałych względów, wybór nazwy zdaje się mieć jednak ogromne znaczenie. Tak na przykład M. Kulicki (1994) poświęca tej sprawie kilka stron swego podręcznika przekonując, że nazwa „poligraf” jest myląca, bowiem „poligraf” kojarzony jest z przemysłem drukarskim, czyli poligrafią.

Nie sądzę, aby były argumenty przekonujące, że dla Polaka nazwa ta jest myląca, a dla Amerykanina, Niemca, Anglika, Czecha czy Rosjanina już nie. We wszystkich tych krajach istnieją też odpowiedniki naszego słowa „poligrafia” w znaczeniu przemysłu drukarskiego. Co więcej, w USA były nawet produkowane maszyny do pisania o nazwie „polygraph” (por. ryc. 1). Największy bodaj słownik języka angielskiego (Oxford English Dictionary, 1971) zna aż 5 znaczeń słowa „polygraph”, 4 znaczenia słowa „polygraphic”, po 3 znaczenia słów „polygrapher” i „polygraphy”, z których większość związana jest z technikami druku.

Wieloznaczność nazw jest typowa dla każdego języka, wystarczy przypomnieć co najmniej trzy zupełnie różne znaczenia w języku polskim słowa „zamek”, dwa różne znaczenia słowa „pokój” itd.

* J. Widacki, *Poligraf czy wariograf?*, „Problemy Kryminalistyki” 2016, 294 (4), s. 15–19.



Ryc.1. Amerykańska maszyna do pisania „Polygraph”

Na liście 325 największych wynalazków i wynalazców wszechczasów, opublikowanej w *Encyclopedia Britannica Almanac*, znajduje się poligraf („polygraph”) i jego konstruktor John A. Larson. Sam prototyp urządzenia znajduje się w Waszyngtonie w Smithsonian Institution. Rzecz ciekawa, Larson swego aparatu nie nazywał „poligrafem”. Używał nazwy „cardio-pneumo-psychograph” (Larson, 1922), a potocznie, slangowo aparat nazywano „sphyggy” (co było skrótem od „sphygmomanometer”). Aparat składał się z klasycznego, znanego od XIX wieku pneumografu, zmodyfikowanego sfigmografu, nazwanego „kardiografem”, pozwalającego na obserwację i zapisywanie nie tylko uderzeń tętna, ale także wahań ciśnienia krwi. Wskazania pneumografu i kardiografu zapisywane były na taśmach zaczerpionego papieru owiniętego na bębnach kimografu. Dodać należy, że pierwszy kimograf skonstruowany został w już 1840 roku, przez niemieckiego fizjologa Karla Ludwiga. Pierwszy sfigmograf skonstruował w 1854 roku inny niemiecki fizjolog, Karl von Vierordt. Powszechnie używane, od końca XIX wieku, pneumografy były konstrukcji francuskiego naukowca i wynalazcy Etienne-Jules Marey’a.

Konstruktor kolejnego, skonstruowanego w latach 20. XX wieku w Berkeley, aparatu do instrumentalnej detekcji kłamstwa, kapitan Clarence D. Lee, nazwał swój aparat „The Berkeley psychograph”. Był to również aparat dwukanałowy (pneumograf plus kardiograf) wyposażony dodatkowo w marker pytań. Reakcje i moment zadania pytania znaczone były pisakami atramentowymi na przesuwającej się taśmie papierowej. Pod tym względem, „psychograf” Lee był podobny do „poligrafu atramentowego” szkockiego kardiologa J. Mackensiego, uważanego mylnie za pierwszy aparat o tej nazwie (por. niżej).

Zanim Leonard Keeler uzupełnił dwukanałowy aparat psychogalwanometrem i nazwał „poligrafem”, używał nazwy „emotograf” (emotograph). Od czasu, kiedy w latach 30. XX wieku Keeler rozpoczął produkcję aparatów o nazwie „Keeler polygraph”, nazwa poligraf stała się powszechnie używaną w USA nazwą aparatów rejestrujących równocześnie przebieg czynności oddychania, zmiany reakcji skórno-galwanicznej, zmiany pracy układu krążenia (częstotliwość tętna i względne wahania ciśnienia krwi), ewentualnie jeszcze inne zmienne fizjologiczne (np. reakcję pletysmograficzną, napięcie mięśni). Równocześnie, cały czas funkcjonowała potoczna nazwa „wykrywacz kłamstwa” (lie-detector).

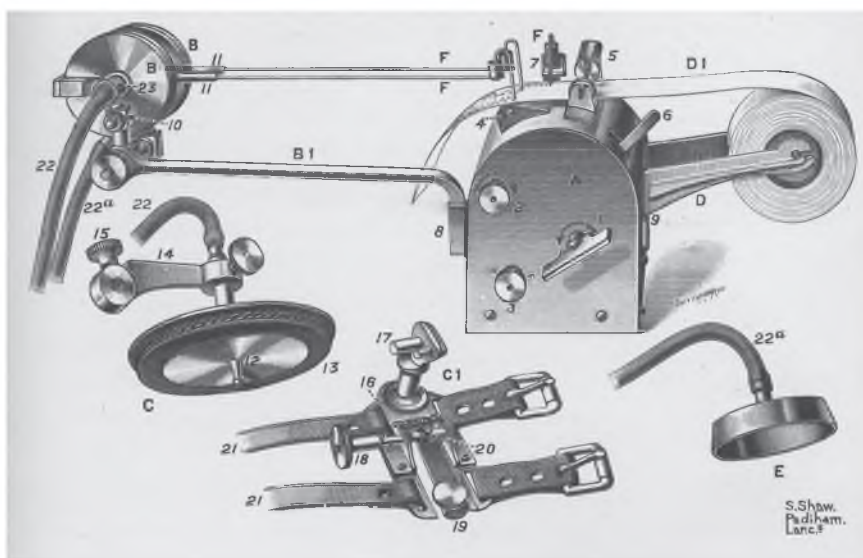
Niezależnie od tego, niektórzy producenci nadawali poszczególnym modelom rozmaite nazwy handlowe. Tak na przykład firma Stoelting jeden z modeli z lat 50. nazwała „deceptografem” („de-

ceptograph”), późniejsze modele, „interografem” („Interograph”), czy „Emotional Stress Monitor”, „Polyscribe”.

Zdaniem Freda Inbaua pierwszym, który użył nazwy „poligraf” („polygraph”) dla nazwania aparatu rejestrującego równocześnie wiele funkcji organizmu, był wspomniany wyżej szkocki kardiolog James Mackenzie (Inbau, 1953). Konstrukcję Mackenzie’go Inbau wprost nazwał „the first polygraph” („pierwszy poligraf”).



Ryc. 2. Poligraf Keelera



Ryc. 3. Poligraf Mackenziego

Mackenzie rzeczywiście zestaw przez siebie skonstruowany, złożony z kimografu z więcej niż jednym urządzeniem odnotowującym jakąś funkcję organizmu, nazywał „poligrafem”. Tak na przykład jego „poligraf kliniczny” („clinical polygraph”) rejestrował na tej samej taśmie zapisy tętna („radial pulse”), uderzenia koniuszkowe serca („apex beat”), tętno z tętnicy szyjnej („carotid pulse”), tętno żyłne („venous”) oraz ruchy oddechowe (respiratory movements”) (Mackenzie, 1910). Kolejną inno-

wacją było zbudowanie „poligrafu atramentowego” („the ink polygraph”), który wykreślał zapisy kilku funkcji organizmu przy pomocy atramentowych pisaków na papierowej taśmie. Dotychczas zapisy wykonywane były albo suchym rysikiem na zaczernionym (lub rzadziej: nawoskowanym) papierze, albo przez cień rzucany na światłoczuły papier. Takie atramentowe pisaki stosował później Leonard Keeler w konstruowanych przez siebie poligrafach. Natomiast zapisy na światłoczułym papierze wykonywał „fotopoligraf” Darrow’a, sprowadzony do Polski w latach 30. XX wieku i wykorzystywany w badaniach emocji wykonywanych w Państwowym Instytucie Higieny Psychiczej w Warszawie.

Nie negując znaczenia udoskonaleń, jakie w sprzęcie zapisującym reakcje organizmu poczynił Mackenzie, stwierdzić trzeba, że wbrew temu, co sądził Inbau, a następnie dziesiątki autorów na całym świecie, powołujących się na Inbaua, Mackenzie nie wymyślił nazwy „poligraf” na określenie aparatu rejestrującego symultanicznie więcej niż jedną funkcję organizmu.

Nazwa ta funkcjonowała w nauce europejskiej, w tym znaczeniu, na długo przed publikacjami Mackenziego. Także w języku polskim. Napoleon Cybulski, w *Fizjologii człowieka* wydanej w 1895, dla określenia urządzenia sprzężonego z kimografem, zapisującego równocześnie więcej niż jedną funkcję organizmu, używa nazwy **poligraf**. Identycznie urządzenie także **poligrafem** nazywa w dwadzieścia lat później Zbyszewski (1914). Jak widać nazwa „poligraf”, dla urządzenia wykonującego równocześnie zapisy więcej niż jednej funkcji organizmu, nie jest, jak uważają niektórzy, kalką z języka angielskiego, ale była w tym znaczeniu używana w nauce polskiej znacznie wcześniej, niż znalazła się w tym znaczeniu w języku angielskim.



Ryc. 4. Kimografy ze zbiorów UJ

Jeśli idzie o nazwy aparatu, zwanego potocznie „wykrywaczem kłamstwa”, to podobnie, jak w przypadku innych urządzeń technicznych, nazwy nadawali albo konstruktorzy, albo producenci. Nie trzeba przypominać, że autor nazwy „wariograf”, prof. Paweł Horoszowski, nie należał ani do jednych, ani do drugich. Najpierw zwalczał poligraf, nazywając go „urządzeniem przyczyniającym się do pogłębiania atmosfery zastraszania osoby przesłuchiwanej” (Horoszowski, 1958), następnie, już w kilka lat później, gdy sam taki aparat zakupił nie mając do tego żadnego przygotowania, wykorzystał go w sprawach karnych, w sposób urągający elementarnym zasadom prowadzenia takich badań, obciążając swą opinią oskarżonego. Opis i ocena badań, wykonanych przez Horoszowskiego w 1963 i 1964 roku, są znane z literatury (Horoszowski, 1965; Krzyścin, 1977; Widacki, 1981, 2014). Liczba błędów popełnionych w tych badaniach, wytkniętych Horoszowskiemu w literaturze, jest długa i nie miejsce tu, by ja powtarzać. Dość może przypomnieć jeden z tych błędów, bardziej niż kardynalny. Mianowicie oznajmijając badanemu, że badanie jest już zakończone, zadał mu zniechęcające pytanie

krytyczne i oczywiście odnotował reakcję. Najbardziej elementarna wiedza dotycząca badań poligraficznych pokazuje, że w takiej sytuacji, każdy badany zareagowałby na pytanie krytyczne, o dowolnej treści, z pytaniem o udział w zabójstwie Kennedy'ego czy Cesara włącznie. Ta oczywista reakcja, była dla Horoszowskiego dowodem na to, że badany na tak zadane pytanie kłamie, a zatem jest winien zarzucanego mu czynu.

Można mieć wątpliwości, czy Horoszowski, nie będąc ani konstruktorem aparatu, ani jego producentem, przy wątpliwych zasługach we wprowadzaniu poligrafu do praktyki śledczej w Polsce, miał zasługujące na aprobatę prawo do wprowadzania swojej nazwy urządzenia. Ja takie wątpliwości mam i dlatego nazwy wariograf nie używam. Używam starej, polskiej nazwy „poligraf”, mając w dodatku satysfakcję, że jest ona powszechnie używana w świecie.

Mogę zrozumieć, że ktoś takich wątpliwości nie ma, albo z jakiegoś powodu przywiązał się do nazwy wariograf i dlatego jej używa. Niechże więc teraz czyni to świadomie.

Piśmiennictwo

- Cybulski N., *Fizjologia człowieka*, wyd. 2, t. I, część I i II, Kraków 1895.
- Horoszowski P., *Eksperymentalno-testowa metoda wariograficzna w śledczej i sądowej ekspertyzie psychologicznej*, „Przegląd Psychologiczny” 1965, 9.
- Horoszowski P., *Kryminalistyka*, PWN, 1958, Warszawa.
- Horoszowski P., *Od zbrodni do kary*, Bibl. Problemów, 1963, Warszawa.
- Ibau E., *The first polygraph*, „Journal of Criminal Law, Criminology and Police Science” 1953, 43, s. 679.
- Kreuz M., *Podstawy psychologii: studium nad metodami i pojęciami współczesnej psychologii*, „Czytelnik”, 1949, Warszawa.
- Krzyścin A., *Eksperci z własnej nominacji*, „Gazeta Sądowa i Penitencjarna” 1977, 4, s. 7.
- Kulicki M., *Kryminalistyka. Wybrane problemy teorii i praktyki śledczo-sądowej*, Wydawnictwo Naukowe Uniwersytetu Mikołaja Kopernika, 1994, Toruń.
- Larson J.A., *The cardio-pneumo-psychogram and its use in the study of emotions with practical applications*, „Journal of Experimental Psychology” 1922, 5 (5), s. 323.
- Mackenzie J., *Diseases of the heart*, Oxford University Press, 1910, London.
- The Compact Edition of the Oxford English Dictionary. Complete text reproduced micrographically*, vol. II, P-Z, Oxford University Press, 1971, s. 1090.
- Widacki J., *Wprowadzenie do problematyki badań poligraficznych*, Wyd. MSW, 1981, Warszawa.
- Widacki J. (ed.), *Badania poligraficzne w Polsce*, Oficyna Wydawnicza AFM, 2014, Kraków.
- Zbyszewski L., *Fizjologia i oddychanie. Mechanizm i unerwienie*, Drukarnia Uniwersytetu Jagiellońskiego, 1914, Kraków.

Streszczenie

W Polsce, i tylko w Polsce, poligraf nazywany bywa też „wariografem”. Argumentem, przemawiającym zdaniem niektórych za nazwą „wariograf” jest to, że nazwa „poligraf” rzekomo jest myląca i kojarzy się z „poligrafia”, czyli drukarstwem. Autor przekonuje, że w innych krajach dwuznaczność nazwy poligraf nie powoduje dezorientacji. Podaje wiele nazw synonimicznych do nazwy „poligraf”, podaje też jak nazywano w Ameryce aparat, potocznie zwany wykrywaczem kłamstwa (lie-detector), zanim przyjęła się ostatecznie nazwa „poligraf” (polygraph). Pokazuje na koniec, że nazwa „poligraf” używana była w polskiej literaturze fizjologicznej jeszcze w wieku XIX na oznaczenie urządzenia rejestrującego równocześnie więcej, niż jedną funkcję fizjologiczną organizmu ludzkiego

Słowa kluczowe: poligraf, wariograf, historia poligrafu

Jan Widacki

Polygraph or Variograph?*

Summary

In Poland and only in Poland, a polygraph is occasionally called a “variograph”. For some, the argument in favor of the name “variograph” is that the term “polygraph” is allegedly misleading as it can be associated with the printing industry. The author argues that in other countries, the ambiguity of the name “polygraph” does not cause confusion. Furthermore, the author mentions several names synonymous with “polygraph” and recalls how the name of the device commonly known as the “lie-detector” has evolved in the US before it eventually became known as the “polygraph”. Finally, the author proves that the name “polygraph” was in use in the Polish physiological literature already in the 19th century, for denominating the device capable of simultaneously registering more than one physiological function of the human body.

Key words: polygraph, variograph, history of polygraph

The name “variograph” was coined by Prof. Paul Horoszowski, who used it for the first time in his expertise issued on June 27, 1963, in the case of manslaughter run by the District (formerly Voivodeship) Prosecutor’s Office in Olsztyn (file no. liDs. 25/63). In 1965, Prof. Horoszowski introduced this term into the scientific literature (Horoszowski, 1965), and two years earlier used it in his popular science book entitled “*Od zbrodni do kary*” [From Crime to Punishment] (Horoszowski, 1963).

In fact, the name was first formulated by the prosecutor in a notice of appointment of an expert, ordering the expert to “conduct variograph analysis”. However, it was certainly not the prosecutor who coined this term, and the fact that he used it in the notice of appointment indicates that he had it previously consulted with the expert to be appointed.

Earlier, for example in his 1958 book entitled “*Kryminalistyka*” [Criminalistics], Horoszowski used the American name “lie-detector”. As a curiosity, an earlier attempt can be mentioned to introduce the name “odkłamywacz” [debunker] into Polish literature (Kreuz, 1949), however, with no success.

In Polish terminology, both professional and scientific, the term “variograph” is used interchangeably with “polygraph”. The latter is used worldwide in a variety of linguistic variants (e.g. “polygraph” in English, “der Polygraf” in German, “poligraf” in Russian). The name “variograph” is used exclusively in Poland.

Both names are basically synonymous. “Polygraph” comes from the Greek *poloi* – much and, *graphos* – writing. In turn, in the term “variograph”, the first Greek segment was replaced with “vario” – a word derived from Latin (from *varius* – various), while the second Greek segment was left unchanged. Incidentally, the Polish word “wariat” [madman] (“different”, “different from normal”), has the same etymology.

In the author’s opinion, there is nothing wrong with using the name “variograph”. Moreover, the author has repeatedly stressed that it is not important whether the analysis is referred to as “polygraphic”

* J. Widacki, *Polygraph or Variograph?*, „Problemy Kryminalistyki” 2016, 294 (4), pp. 88–92.

or “viariographic”, as long as it is performed in accordance with the rules applicable (Widacki, 2014). However, for some authors, the choice of name seems, for unexplainable reasons, to be of the utmost importance. So for example, M. Kulicki (1994) dedicates a few pages of his textbook to convincing the reader that the name “polygraph” is misleading, as it is associated with the printing industry (also referred to as polygraphy).

The author does not consider the opinion that the name is misleading for the Poles, but not for the Americans, Germans, English, Czechs or Russians, a compelling argument. In each of these countries there is an equivalent of the word “polygraphy” in the meaning of the printing industry. What’s more, in the United States a typewriter has been even manufactured, called the “polygraph” (cf. fig. 1). The seemingly largest dictionary of the English language (Oxford English Dictionary, 1971) knows 5 meanings of the word “polygraph”, 4 of the word “polygraphic”, 3 for each of the words “polygrapher” and “polygraphy”, most of which are related to printing techniques.

Name ambiguity is typical of each language. For example, the Polish language knows three different meanings of the word “zamek” [castle, lock, zipper – translator’s note], two of the word the word “pokój” [room, peace – translator’s note], and so on.

The polygraph and its constructor, John A. Larson are included on the list of 325 greatest inventions and inventors of all time, published in the Encyclopedia Britannica Almanac. The prototype of this device is held by the Smithsonian Institution, Washington D.C. Interestingly, Larson did not name his construction “polygraph”, but referred to it as “cardio-pneumo-psychograph” (Larson, 1922). A popular slang term for the polygraph was “sphyggi” (an abbreviation from “sphygmomanometer”). The apparatus consisted of a spiograph (a device known since the 19th century) and a modified sphygmograph, named “cardiograph”, allowing the observation and recording not only the pulse, but also the fluctuations in blood pressure. The indications of the spiograph and cardiograph were recorded on tapes made of blackened paper, wrapped around the reels of a kymograph. It should be added that the first kymograph was designed already in the year 1840, by the German physiologist Karl Ludwig. The first sphygmograph was designed in 1854 by another German physiologist, Karl von Vierordt. Finally, the spiograph, commonly used since the end of the 19th century, was the construction of the French scientist and inventor, Etienne-Jules Marey.

Another lie detection device was designed in the 1920s, in Berkeley, by Captain Clarence D. Lee, who named it “The Berkeley psychograph”. It was also a dual-channel apparatus (spiograph plus cardiograph), additionally equipped with a question marker. The reactions and the moments of asking questions were marked on the moving paper tape by ink plotter pens. In this respect, Lee’s “psychograph” was similar to the “ink polygraph” developed by the Scottish cardiologist, J. Mackenzie, which was erroneously considered the first apparatus with this name (see below).

Prior to complementing his dual-channel apparatus with psychogalvanometer and naming it “polygraph”, Leonard Keeler used the name “emotograph”. Since Keller began manufacturing the apparatus called “Keeler polygraph” in the 1930s, the name “polygraph” has become a term widely used in the US for the devices simultaneously registering the course of respiration, dermal-galvanic reaction, changes in the operation of the cardiovascular system (pulse rate, relative fluctuations in blood pressure), or other physiological variables (for example, plethysmographic reaction, muscle tension). At the same time, a common name “lie detector” has been in parallel use.

Notwithstanding the above, some manufacturers have denominated their individual models with a variety of trade names. Stoelting company, for example, used the term “deceptograph” for one of the models from the 1950s, and “interrograph”, “Emotional Stress Monitor” or “Polyscribe” for subsequent models.

According to Fred Inbau, the name “polygraph”, with reference to the device capable of simultaneously recording multiple bodily functions, was for the first time used by the above mentioned Scottish cardiologist James Mackenzie (Inbau, 1953). Thus Mackenzie’s construction was named “the first polygraph” by Inbau.

Indeed, Mackenzie had actually used the name “polygraph” for his construction made up of a kymograph and more than one device registering a bodily function. For example, his “clinical polygraph” registered on the same tape the “radial pulse”, “apex beat”, the pulse of the carotid artery (“carotid pulse”), “venous pulse” and “respiratory movements” (Mackenzie, 1910). Another innovation was the construction of the “ink polygraph”, which plotted the records of several bodily functions on a paper tape, using ink plotter pens. At that time, the records were either made with a dry stylus on blackened, or, more rarely, waxed paper, or by casting the shadow on light-sensitive paper. Such ink plotter pens were later applied by Leonard Keeler in his polygraph constructions. The recording on light-sensitive paper was applied in Darrow’s “photopolygraph”, a device brought to Poland in the 1930s and used for studying emotions at the National Institute of Mental Hygiene in Warsaw.

Without denying the importance of improvements made by Mackenzie to the device registering bodily reactions, it should be emphasized that, contrary to Inbau’s convictions, shared by dozens of authors citing him worldwide, Mackenzie did not invent the name “polygraph” to be used with reference to the device simultaneously registering more than one bodily function.

It has actually functioned within the European science community long before Mackenzie’s publications. Also, as regards the Polish language, the name polygraph was used by Napoleon Cybulski in his work entitled “*Fizjologia człowieka*” [Human physiology], published in 1895, for the device coupled to the kymograph, capable of simultaneously recording more than one bodily function. Twenty years later, an identical device was assigned the name polygraph by Zbyszewski (1914). As can be seen, the name “polygraph” in the meaning of the device simultaneously registering more than one bodily function is not, as considered by some, merely a loan translation from English, since it was in use within the Polish science community long before it entered the English language.

As regards the terminology concerning the device commonly known as a “lie detector”, just like in the case of other technical devices, the names were coined either by the constructors or manufacturers. Needless to recall, the author of the name “variograph”, Prof. Paweł Horoszowski, did not belong to either group. He initially opposed the polygraph, calling it “a device contributing to enhancing the atmosphere of intimidation of the person being examined” (Horoszowski, 1958). However, a few years later, after purchasing a similar apparatus, Horoszowski used it for the purpose of criminal proceedings, without having a proper background, in a manner that undermined elementary rules for conducting such studies, issuing an opinion that was damning for the accused. The description and assessment of research carried out by Horoszowski in 1963 and 1964, can be found in the literature (Horoszowski, 1965; Krzyścin, 1977; Widacki, 1981, 2014). The list of errors committed during these studies, which fueled criticism among other authors, is lengthy and beyond the scope of this article. It will therefore be sufficient to recall only one of these errors, more than serious. Namely, after informing the person being examined that the study was completed, Horoszowski unexpectedly asked a critical question and noted the response given. Even the most elementary knowledge of polygraphic analyses should be sufficient to realize that, in such a situation, any test subject would react to the critical question, of any content, including about his/her involvement in Kennedy’s or Ceasar’s assassination. Such an obvious reaction was considered by Horoszowski as proof that the test subject responded untruthfully to the question asked, and was therefore guilty of the alleged deed.

One can have doubts whether Horoszowski, being neither the constructor, nor manufacturer of the polygraph, and having questionable contribution to implementing it into investigatory practice in Poland, had a moral right to propagate the name of his choosing for this device. This author shares the above doubts and thus uses an old Polish name “polygraph”, instead of “variograph”, while being satisfied that the same name is commonly accepted worldwide.

The author can understand that certain persons may have no such doubts, or for some reason became attached to the name “variograph”, and thus continue to use it. At least, they will be able to do it knowingly after reading this article.

Bibliography

1. *Badania poligraficzne w Polsce*, (ed.) J. Widacki, AFM Publishing House, Kraków 2014.
2. Cybulski N., *Fizjologia człowieka*, vol. I, part I i II, Kraków 1895, 2nd. Edition.
3. Horoszowski P., *Eksperymentalno-testowa metoda wariograficzna w śledczej i sądowej ekspertyzie psychologicznej*, „Przegląd Psychologiczny” 1965, 9.
4. Horoszowski P., *Kryminalistyka*, PWN, Warszawa 1958.
5. Horoszowski P., *Od zbrodni do kary*, Biblioteka Problemów, Warszawa 1963.
6. Ibaud F., *The first polygraph*, “*Journal of Criminal Law, Criminology and Police Science*” 1953, 43, p.679.
7. Kreuz M., *Podstawy psychologii: studium nad metodami i pojęciami współczesnej psychologii*, Czytelnik, Warszawa 1949.
8. Krzyścin A., *Eksperci z własnej nominacji*, „Gazeta Sądowa i Penitencjarna” 1977,4, p. 7.
9. Kulicki M., *Kryminalistyka. Wybrane problemy teorii i praktyki śledczo-sądowej*, Scientific Nicolaus Copernicus University Publishing House, Toruń 1994.
10. Larson J.A., *The cardio-pneumo-psychogram and its use in the study of emotions with practical applications*, “*Journal of Experimental Psychology*” 1922, 5 (5), p. 323.
11. Mackenzie J., *Diseases of the heart*, Oxford University Press, London 1910.
12. *The Compact Edition of the Oxford English Dictionary. Complete text reproduced micrographically*, vol. II, P-Z, Oxford University Press, Oxford 1971, p. 1090.
13. Widacki J., *Wprowadzenie do problematyki badań poligraficznych*, MSW Publishing House, Warszawa 1981.
14. Zbyszewski L., *Fizjologia i oddychanie. Mechanizm i unerwienie*, Jagiellonian University Printing House, Kraków 1914.

Jan Widacki

The First Polygraph?*

Первый полиграф?

Key words: Polygraph, Inc-Polygraph, History of Lie-Detection

In 1953, *The Journal of Criminal Law, Criminology and Police Science* (No. 5, pp. 679—681) published by the Northwestern University in Chicago reprinted an article by James Mackenzie on “The Ink Polygraph”, originally published by the *British Medical Journal* in 1908 (Vol. 1, p. 1411).

Mackenzie’s article was preceded by an introduction by Professor Fred E. Inbau (1909—98), an unquestioned authority not only on polygraph testing and interrogation techniques, but also on criminal trials and forensic studies (he was among others, a long-term director of the Crime Detection Laboratory). The material as a whole, i.e. Mackenzie’s article together with Inbau’s introduction is entitled “The First Polygraph”.

In his introduction to Mackenzie’s article, Inbau remarked that when the history of lie detection techniques is discussed, hardly anyone notices that the name “polygraph” existed at least as early as 1906 (Inbau 1953, 678). He goes on to comment that the device in question, albeit constructed not for lie detection but for medical examinations, was nonetheless based on the same principles of construction as a lie detector. Following Fred Inbau, many authors repeated this information, as imprecise as it was untrue, suggesting that the device constructed by Mackenzie was the first one to be called “a polygraph” (Kniaziew, Warlamow 2012; Matte 1996).

In his article, Mackenzie wrote that at a meeting of the Medical Section of the British Medical Association in Toronto in 1906, he demonstrated a method of recording the movements of circulation by means of an ink polygraph. The polygraph recorded the circulation, registering pulse frequency and respiratory functions. This was nothing new, however. Since the end of the 19th century, respiration had been analysed with a pneumograph: a device that recorded the movements of the chest or diaphragm. The most popular construction was Marey’s pneumograph, developed by a French scientist and inventor, Étienne-Jules Marey (1830—1904).

Heart activity and blood circulation were observed with a sphygmograph, which made it possible to observe the pulse. The first sphygmograph was constructed in 1854 by a German physiologist, Karl von Vierordt (1818-84). It consisted of a bar placed on an artery that was fastened to the forearm with a strap. The movement of the bar was a function of pulsing of the artery.

A design that could be connected to the pneumograph to produce the respiration curve, or with a sphygmograph to plot a curve recording the action of the heart (pulse beats), known as the kymograph, was constructed by a German physiologist, Karl Ludwig (1816—95). Some authors believe that the kymograph was constructed independent of Ludwig and even earlier, namely in 1846, by an Italian, Carlo Matteucci (Emeryk-Szajewska 2008). Its construction was relatively simple: a clockwork was attached to a drum and it was wrapped in a sheet of blackened paper revolving at a constant speed. The device, whether a sphygmograph or a pletysmograph, transmitted mechanical impulses collected

* J. Widacki, *The First Polygraph?*, “European Polygraph” 2016, Vol. 10, No. 3(37), s. 111–116, DOI 10.1515/ep-2016-0014. Project DEC-2013/11/B/HS5/03856 funded by National Science Centre.

by the original device to the stylus of the kymograph, which drew (or more precisely: erased), a white curve on the blackened sheet.

The sphygmograph, sometimes also known as a cardiograph and combined with a kymograph, was used late in the 19th century in experiments conducted by Angelo Mosso (Mosso 1891) Cesare Lombroso (Lombroso 1891) and others. A pneumograph connected to a kymograph was used in the experiments of both Benussi (benussi 1914) and Abramowski (Abramowski 1913)

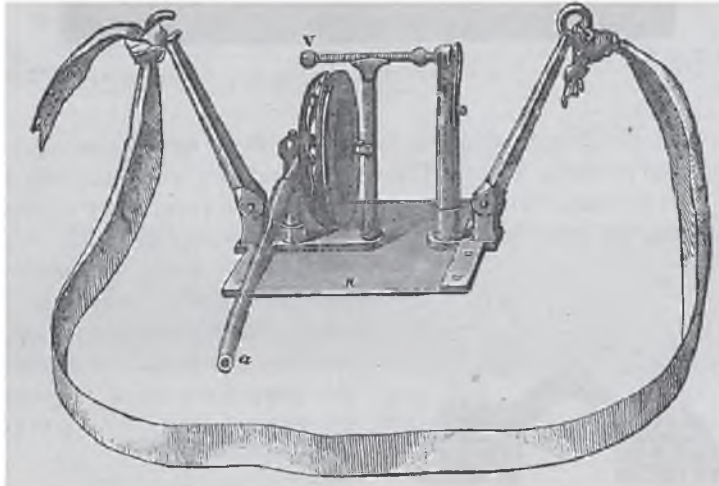


Fig. 1. Marey's pneumograph (N. Cybulski 1895).

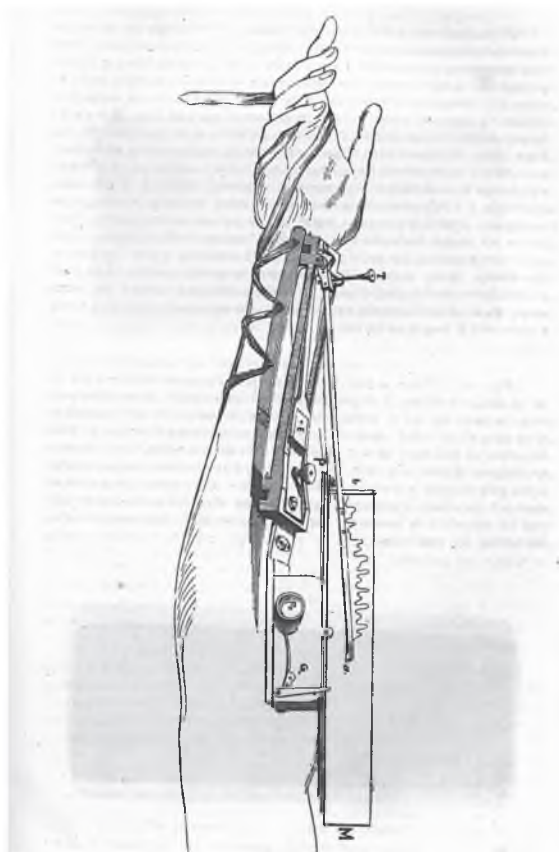


Fig. 2. Von Vierort's sphygmograph (H. Hoyer: Fiziologia, Warszawa 1872).

The kymograph was used as early as the 19th century for the simultaneous recording of data coming from the pneumograph and sphygmograph. The device thus constructed was known as a “polygraph”, i.e. a device for recording more than one function (Greek: *poloi* — multiple, *graphos* — to write). Such a name was used to denote the combination of a kymograph with a sphygmograph or pneumograph in the 19th century by Napoleon Cybulski (Cybulski 1895) and somewhat later by Leon Zbyszewski (Zbyszewski 1914). Thus European continental literature had applied the name “polygraph” for at least over a decade before Mackenzie used it. This demonstrates that rather than inventing the name, Mackenzie just mentioned it. The range of devices he mentioned included “sphygmograph”, “the polygraph”, “the clinical polygraph”, and “the ink polygraph”.

According to Mackenzie, the clinical polygraph “can be used for taking, at the same time and on the same recording surface, tracings of the radial pulse, with tracings of the apex beat, carotid, venous, or liver pulse, or respiratory movements, and its size permits it being carried out with the greatest facility, and readily employed in general practice” (Mackenzie 1910, 68). He explains further that “the essential part of the instrument [clinical polygraph] is a small cup for receiving the impressions of the pulsations, a tube for transmitting the impressions to a tambour and the lever the tambour is attached to (...) sphygmograph” (Mackenzie 1910, 69).

The “ink polygraph” was an enhancement of the “clinical polygraph”. Mackenzie arrived at the conclusion that “it was not convenient when the movements had to be recorded over a long period, as where the irregularities were infrequent, or where they varied, or where respiratory curves were required. [He] therefore conceived the idea of constructing an instrument that would take tracings of an indefinite length, where the employment of ink would enable a roll of paper to be unwound, and save as well the inconvenience of blackening and varnishing” (Mackenzie 1910, 72).

Subsequently, with his assistant S. Shaw, Mackenzie constructed a device called the “ink polygraph”. The novelty was that rather than turning a drum wrapped in a sheet of paper, as in kymographs, the clockwork (A) set in motion a roll of paper (D). Another clockwork moved the time-marking pen (F), two tambours (BB) and their levers (FF). The levers were connected to pens, and ink was poured into a small container and it reached the pens through capillaries.

There were two fundamental changes in the improvement introduced by Mackenzie. First, he replaced the tambour wrapped in blackened paper used in kymographs with a roll of paper. The stylus or slate pencil erasing a white curve on the blackened paper was replaced by an ink pen that drew a curve on the white roll. This was certainly a major development, and the upgraded system of recording was later used in analog polygraphs.

Thus, James Mackenzie certainly modernised the manner of recording physiological functions of the human organism that had been observed and recorded many years before him. It should likewise be recognised that he was not the first to use the name “polygraph” to denote a device that simultaneously records more than one single physiological function of the human organism as the name had been used over a decade before him for the combination of a sphygmograph and pneumograph with a kymograph.

References

- Abramowski E. (1913): *Oddech jako czynnik życia duchowego. Wzruszenie i wola.*, *Prace z Psychologii Doświadczalnej*, 1913, I, pp. 77–162 [in Polish].
- Benussi V. (1914): *Die Antmungssymptome der Lüge*, *Archiv ges. Psychol.*, 1914, 31, pp. 244–273.
- Cybulski N. (1895): *Fizjologia człowieka*, 2nd ed. Vol. I, parts I and II, Kraków 1895, pp. 430–431 [in Polish].
- Emeryk-Szajewska B. (2008): *Krótką historia powstania i rozwoju elektromiografii i elektroneurologii* [in:] *Neurologia kliniczna*, Vol. I, B. Emeryk-Szajewska, M. Niewiadomska-Wolska (eds.), *Medycyna Praktyczna*, Kraków 2008, pp. 47–54.
- Kniaziev V., Varlamov G. (2012): *Poligraf i ego praktičeskoje primienienie*, Moskwa 2012, p. 55.

- Lombroso C. (1891): Człowiek zbrodniarz [Polish translation], Warszawa 1891–92.
- Mosso A. (1891): Strach. Studium popularnonaukowe [Polish translation], Warszawa 1891, pp. 95 and ff.
- Mackenzie J. (1908): The ink polygraph, British Medical Journal 1908, 1, 1411.
- Mackenzie J. (1910): Diseases of the heart, 2an ed. Oxford University Press, London 1910.
- Matte J. (1996): Forensic psychophysiology. Using the polygraph. Scientific truth verification-lie detection, J.A. M. Publications, Williamsville (N.Y.) 1996, p. 19.
- Zbyszewski L. (1914): Fizyologia oddychania. Mechanizm i unerwienie, Drukarnia Uniwersytetu Jagiellońskiego, Kraków 1914, p. 22.

Jan Widacki, Romuald Polczyk, Marcin Gołaszewski, Michał Widacki, Paweł Miazga,
Karolina Dukała, Anna Szuba-Boroń, Marcin Tarabula, Aleksandra Kuhny

Część II

Eksperyment

1. Cel eksperymentu

Celem eksperymentu było uzyskanie do dalszych analiz materiału z badań poligraficznych, którym poddane były osoby „niewinne” (w badaniu prawdomówne) oraz „winne” (w badaniu kłamiące). Było to typowe badanie „laboratoryjne”, w którym osoby „winne” uczestniczyły w zaaranżowanym zdarzeniu.

Zebrano grupę 39 ochotników, studentów i studentek Wydziału Prawa, Administracji i Stosunków Międzynarodowych oraz Wydziału Nauk o Bezpieczeństwie Krakowskiej Akademii im. Andrzeja Frycza Modrzewskiego. Byli to studenci studiów stacjonarnych i niestacjonarnych. Tak więc wszyscy badani mieli wykształcenie średnie i niepełne wyższe. W tej sytuacji można założyć, że wszyscy dysponowali wystarczającym dla badań poligraficznych stopniem rozwoju umysłowego¹.

Nikt z badanych nie deklarował żadnych schorzeń (w tym psychicznych), które mogłyby być przeciwwskazaniem do skutecznych badań poligraficznych.

Z uwagi na to, że studenci sami zgłaszali się do badań, można podejrzewać, że osoby o ekstrawersji wyższej niż przeciętna były w tej grupie nadreprezentowane. Jednak wedle aktualnego stanu wiedzy nie powinno to mieć znaczenia dla skuteczności badań poligraficznych, którym zostali poddani. Pod względem poziomu wybranych cech osobowości średnia w próbie nie odbiegała od średniej w populacji ($p < 0,05$).

W grupie badanych było 13 mężczyzn (33,3%) i 26 kobiet (66,7%), co mniej więcej odpowiada strukturze populacji studentów wspomnianych wydziałów, gdzie liczebnie dominują kobiety. Rozpiętość wieku osób badanych była znaczna (od 20 do 43 lat), przy czym liczebnie dominowali ludzie młodzi. W wieku od 20 do 22 lat było 25 osób badanych (64% badanej grupy). Średnia wieku w całej grupie wynosiła 22,1 (SD = 4,89). Zatem w badanej grupie nie było ani osób niepełnoletnich, ani osób w wieku podeszłym.

¹ Jak wiadomo, do badań poligraficznych nie nadają się z zasady osobnicy o ilorazie inteligencji < 80 (to jest poniżej granicy ociężałości umysłowej); por. *Kryminalistyka*, red. J. Widacki, Kraków 2016, s. 423.

Wszystkie osoby, które zgłosiły się do udziału w eksperymencie, były uprzednio informowane, że będą badane psychologicznie, a następnie poligraficznie. Wszystkie wyraziły zgodę na poddanie się takim badaniom.

Pierwszym etapem były badania psychologiczne. Do badań wykorzystano 3 narzędzia:

- EPQ-R – Kwestionariusz Osobowości Eysencka (w polskiej adaptacji Brzozowskiego i Drwała 1995) (pierwsze 100 pozycji);
- FCZ-KT – Formalna Charakterystyka Zachowania (Cyniak-Cieciura, Zawadzki, Strelau, 2016);
- NEO-FFI, tzw. Wielka Piątka, Inwentarz Osobowości NEO-FFI (Costa, McCrae, 1992, w polskiej adaptacji Zawadzki, Strelau, Szczepaniak, Śliwińska, 2007) (ostatnie 60 pozycji).

Narzędzia te pozwoliły w badanej grupie zdiagnozować:

- 1) ekstrawersję, neurotyzm, psychotyzm, oraz skłonność do przedstawiania się w lepszym świetle (EPQ-R);
- 2) temperament, opisywany w kategoriach: żwawość, perseweratywność, wrażliwość sensoryczną, reaktywność emocjonalną, wytrzymałość, aktywność (FCZ-KT);
- 3) pomiar osobowości w skalach: ekstrawersja, ugodowość, sumienność, neurotyczność, otwartość na doświadczenie.

Zakładaliśmy, że te cechy mogą mieć pewne znaczenie przy badaniach poligraficznych.

Osoby badane nie różniły się ww. cechami osobowości pomiędzy warunkami eksperymentalnymi (winni vs. niewinni, wszystkie $p < 0,05$). Seria niezależnych analiz regresji pokazała, że żadna z cech testowanych osobowości nie ma wpływu na 1) trafność ocen poligraferów, 2) czułość (trafność oceny osób kłamiących) ani 3) swoistość (trafność oceny osób niekłamiących). Szczegółowe wyniki przedstawia tabela 1.

Tabela 1. Seria niezależnych analiz regresji pokazująca wpływ cech osobowości na różne miary trafności ocen poligraferów: trafność ogólna, czułość i swoistość. p oznacza istotność statystyczną, B oznacza niestandardyzowany współczynnik regresji liniowej

	Wpływ na trafność		Wpływ na czułość		Wpływ na swoistość	
	p	B	p	B	p	B
Ekstrawersja (EPQ-R)	0,721	0,007	0,550	0,023	0,672	0,009
Neurotyzm	0,714	0,005	0,754	-0,085	0,847	-0,003
Psychotyzm	0,157	-0,026	0,195	-0,042	0,415	-0,016
Kłamstwo	0,471	0,014	0,406	0,031	0,324	0,019
Wrażliwość emocjonalna	0,215	0,026	0,406	0,031	0,272	0,026
Reaktywność emocjonalna	0,456	0,011	0,711	-0,014	0,530	0,009
Perseweratywność	0,203	0,022	0,642	0,018	0,442	0,014
Aktywność	0,767	-0,005	0,780	0,009	0,681	-0,007
Żwawość	0,311	-0,024	0,843	-0,008	0,663	-0,013
Wytrzymałość	0,599	-0,008	0,203	-0,041	0,393	0,013
Neurotyzm	0,534	0,005	0,800	-0,004	0,723	0,003
Ekstrawersja (NEO-FFI)	0,261	-0,011	0,096	-0,038	0,987	0,001
Otwartość	0,674	0,005	0,920	-0,002	0,390	0,010
Ugodowość	0,632	0,005	0,889	0,002	0,847	0,002
Sumienność	0,749	-0,003	0,741	-0,006	0,372	0,009

2. Osoby badane – podział na role

Ogół badanych podzielono losowo na dwie grupy, przypisując jednej z nich rolę „sprawców”, a drugiej – „niewinnie podejrzanych”. 15 osób pełniło rolę sprawców, a 24 – niewinnych podejrzanych.

„Sprawcy” w towarzystwie jednego z organizatorów eksperymentu udawali się na krytą strzelnicę i dostawali pistolet (atrapa pistoletu glock, strzelająca plastikowymi kulkami, wyrzucanymi energią sprężonego gazu). Ich zadaniem było przeładowanie pistoletu i oddanie trzech strzałów do sylwetki młodej kobiety na kolorowym plakacie. Po wykonaniu tego zadania, uczestnicy eksperymentu byli informowani, że mają zachować w tajemnicy fakt, że byli na strzelnicy i strzelali. W szczególności fakt ten mieli zachować w tajemnicy przed osobą, która będzie ich badała na poligrafie. Mieli więc w czasie tego badania, odpowiadając na pytania krytyczne, kłamać. Byli poinstruowani, by w czasie badania poligraficznego zachowywali się tak, jak wyobrażają sobie, że zachowywałby się sprawca przestępstwa, który nie przyznaje się do winy. Dla wzmocnienia ich motywacji do zachowania w czasie badania poligraficznego podobnego do zachowania sprawcy przestępstwa, dawano im banknot 50-złotowy (równowartość ok. 14 USD). W razie pomyślnego dla nich przejścia przez badanie poligraficzne, to znaczy w sytuacji, kiedy badający ekspert nie wykryłby ich kłamstwa, banknot stawałby się ich własnością. W sytuacji, gdyby ekspert po badaniu poligraficznym, wykrył, że byli na strzelnicy i strzelali, a teraz pytani o to kłamią, musieliby zwrócić banknot.

„Niewinni” nie byli na strzelnicy, nie strzelali, nie mieli pojęcia o tym, co robili „sprawcy”. W czasie badania poligraficznego, na pytania testów mieli odpowiadać zgodnie z prawdą. Byli poinstruowani, że mają się w czasie badania zachowywać tak, jak wyobrażają sobie, że zachowywałaby się osoba niewinnie podejrzana. Dla wzmocnienia ich motywacji, przed badaniem dostawali podobnie jak „sprawcy” kwotę 50 zł, z tym że mieli ją zwrócić, w przypadku, kiedy ekspert pomyliłby się i wskazał, że byli na strzelnicy i strzelali. Tłumaczono im to tak, że w prawdziwym badaniu poligraficznym osoba niewinnie podejrzana ponosi negatywne konsekwencje pomyłki eksperta.

Z każdą osobą, która miała być badana poligraficznie, na krótko przed badaniem osoba organizująca eksperyment (nie przeprowadzająca badanie!) prowadziła „rozmowę motywującą”. Osobom, które wylosowały rolę „niewinnie podejrzanych”, nie strzelały na strzelnicy, nie wiedziały, kto strzelał, nie znały żadnych szczegółów związanych ze strzelaniem), mówiono tak:

Uczestniczysz w ważnym eksperymencie w ramach projektu badawczego finansowanego przez Narodowe Centrum Nauki. Tym samym, masz swój osobisty wkład w rozwój nauki. To nie eufemizm. To prawda. Także od Twojej postawy w tym eksperymencie zależeć będą rzetelne wyniki badań. Stąd nasza prośba. Zastosuj się do podanych zaleceń:

1) w eksperymencie przypadła Ci rola osoby niewinnie podejrzanej. Nie mów nikomu, a szczególnie osobie prowadzącej badanie poligraficzne, że taka rola właśnie Tobie przypadła. Masz do swej niewinności przekonywać osobę, która będzie Cie badać na poligrafie. On nie będzie wiedział, czy przekonując go o tym, mówisz prawdę. Z nikim innym o tej swojej roli w eksperymencie nie rozmawiaj;

2) Ekspert-poligrafer będzie Cię pytał o różne rzeczy. Nie będą to oczywiście pytania naruszające Twoją prywatność. Odpowiadaj mu zgodnie z prawdą. Nic nie wiesz o zdarzeniu na strzelnicy, o udział w którym będzie Cię pytał. On naprawdę nie wie, jaką rolę mają poszczególni badani w eksperymencie, kto jest „sprawcą”, a kto „osobą niewinną”;

3) Staraj się wczuć w rolę osoby niewinnie podejrzanej. Zachowuj się tak, jak zachowywałaby się osoba niewinnie podejrzana, której prawdopodobność ktoś chce sprawdzić badaniem poligraficznym. Tak jak osoba niewinna, na pytania testu odpowiadaj zgodnie z prawdą;

4) Tak jak osoba niewinna w prawdziwym procesie, jesteś w tym badaniu zainteresowany/zainteresowana, aby badający Cię ekspert się nie pomylił, ale przeciwnie, jesteś zainteresowany/zainteresowana, aby potwierdził Twoją prawdomówność. Nie utrudniaj mu zatem w żaden sposób badania;

5) Podobnie jak osoba niewinna poddana badaniom poligraficznym masz coś do stracenia. Błąd eksperta obciąża w postępowaniu dowodowym osobę niewinną. Tutaj, błędne uznanie Cię za osobę kłamiącą też pociągnie za sobą ujemne dla Ciebie konsekwencje. Jeśli ekspert się pomyli, uzna błędnie, że w czasie badania kłamiesz, odpowiadając na pytania krytyczne – będziesz musiał/musiała oddać te 50 zł, które teraz Ci dają. Myślisz, że to nie Twoja wina, że on się pomylił? No nie Twoja. Ale co z tego? Jak w badaniu w autentycznej sprawie ekspert się pomyli i uzna, że osoba prawdomówna i niewinna kłamie, to też nie będzie to jej wina, ale poniesie dużo dalej idące konsekwencje. Jak widzisz, to badanie ma się między innymi przyczynić do tego, by eksperci-poligraferzy nie robili takich pomyłek;

6) Zrób zatem wszystko, aby biegły poligrafer, badając Cię, nie pomylił się. Stosuj się w czasie badania do wszystkich jego instrukcji. Wczuj się w rolę osoby niewinnie podejrzaną. Pamiętaj, że w tego rodzaju badaniach trafny wynik jest osiągalny w 85–90% przypadków. Masz zatem ogromną szansę, że nie będziesz błędnie wskazany/wskazana jako osoba kłamiąca i zatrzymasz w nagrodę te 50 zł.

Z kolei osoby, które w eksperymencie wylosowały rolę „sprawcy”, strzelały na strzelnicy do wizerunku kobiety na afiszu, bezpośrednio przed badaniem otrzymywały następującą instrukcję:

Uczestniczysz w ważnym eksperymencie w ramach projektu badawczego finansowanego przez Narodowe Centrum Nauki. Uczestnicząc w tym eksperymencie – masz swój osobisty wkład w rozwój nauki. To nie eufemizm. To prawda. Także od twojej postawy w tym eksperymencie zależeć będą rzetelne wyniki badań. Stąd nasza prośba: zastosuj się do następujących poleceń:

1) W eksperymencie wylosowałeś/wylosowałaś rolę sprawcy. Nie mów nikomu, że taką rolę wylosowałeś/wylosowałaś. Nie mów nikomu, że byłeś/byłaś na strzelnicy, że strzelałeś/strzelałaś. Gdy trafisz na badanie poligraficzne, nie wolno Ci się do tego przyznać. Masz uparcie, konsekwentnie twierdzić, że na żadnej strzelnicy nie byłeś/byłaś, nie strzelałeś/strzelałaś, nie wiesz, czy i kto tam był i co robił, a zwłaszcza czy strzelał, czy nie. W tej części eksperymentu Twoim podstawowym zadaniem jest zataić przed badającym cię na poligrafie ekspertem, że byłeś/byłaś na strzelnicy, że strzelałeś/strzelałaś;

2) Ekspert-poligrafer będzie Cię pytał o różne rzeczy. Na pewno nie będą to pytania naruszające Twoją prywatność. Pamiętaj, że jego zadaniem i jego celem jest ustalenie, która z badanych przez niego osób jest „sprawcą”, która jest „niewinna”. **Zachowuj się tak, jak wyobrażasz sobie, że zachowywałby się sprawca przestępstwa, który do jego popełnienia się nie przyznaje i został poddany badaniom poligraficznym;**

3) Tak jak w prawdziwej sprawie karnej nieprzyznający się do popełnienia przestępstwa sprawca poddany badaniu poligraficznemu jest zainteresowany tym, by ekspert nie wykrył jego kłamstwa, bo pociągnęłoby to za sobą oczywiste, negatywne dla niego konsekwencje, tak i Ty, jeśli ekspert-poligrafer wykryje Twoje kłamstwo na pytania krytyczne, poniesiesz negatywne konsekwencje tego faktu: będziesz musiał/musiała zwrócić 50 zł, które teraz Ci dają;

4) Mam jednak nadzieję, że uda Ci się przejść przez to badanie pomyślnie, że uda Ci się skutecznie zataić fakt bycia na strzelnicy i strzelania. Tylko nie próbuj tego robić, zaburzając sztucznie zapis poligrafu. Ekspert od razu to wykryje i będzie wiedział, że próbujesz go oszukać. Jeśli uda Ci się pomyślnie dla Ciebie przejść przez badanie, uda Ci się skutecznie zataić fakt bycia na strzelnicy i strzelania, możesz być z siebie dumny/dumna! To udaje się tylko 5–15% badanych. Jeśli będziesz się do nich zaliczać, należą Ci się gratulacje. Masz kwalifikacje na oficera służb specjalnych. Oprócz gratulacji, dostaniesz nagrodę 50 zł, które będziesz mógł zatrzymać i wykorzystać, sprawiając sobie jakąś małą przyjemność.

3. Eksperci-poligraferzy

W eksperymencie badania poligraficzne przeprowadzało trzech kwalifikowanych ekspertów zatrudnionych w instytucji państwowej i wykonujących badania poligraficzne na jej rzecz. Wszyscy mieli podobne przygotowanie zawodowe i podobne doświadczenie w badaniach poligraficznych;

4. Aparat

Do badań używano komputerowego poligrafu produkcji USA, Lafayette LX 4000, rejestrującego: przebieg czynności oddychania (za pomocą dwóch pneumografów zakładanych na klatkę piersiową i przeponę badanego), przebieg pracy serca (tętno i wahania ciśnienia krwi), reakcję skórno-dermalną (odruch skórno-galwaniczny RSD, GSR), przy zastosowaniu czujnika napięcia mięśni.



Ryc. 1. Poligraf Lafayette LX 4000

5. Pokój do badań

Badania odbywały się w specjalnym pokoju w Krakowskiej Akademii, spełniającym wymagania niezbędne dla badań poligraficznych. W szczególności pokój ten izolował badanego od wpływu zewnętrznych bodźców (wzrokowych, słuchowych) które mogłyby zakłócić przebieg badania.



Ryc. 2. Pokój, w którym przeprowadzono badanie.

6. Technika badania

Po dyskusjach na temat wyboru techniki badania, w badaniach wykorzystano rutynowo stosowaną w USA, a także w Polsce technikę Utah Zone Comparison Technique (UZCT). Świadomie zrezygnowaliśmy ze stosowania w badaniach techniki Lykken².

Braliśmy pod uwagę to, że w technikach pytań kontrolnych stosowanych w badaniach eksperymentalnych, w warunkach laboratoryjnych waga pytania kontrolnego dotyczącego konkretnej sytuacji życiowej może z natury rzeczy być większa, niż waga pytania krytycznego dotyczącego wyreżyserowanej, sztucznej sytuacji³. Z drugiej jednak strony, znane są badania eksperymentalne, w których z powodzeniem stosowano techniki pytań kontrolnych. Tak na przykład Barland⁴ w podobnym do naszego eksperymencie stosował technikę Backstera. Dokonując bardzo ostrożnej oceny wykresów, co wyrażało się wysokim procentem wyników nierozstrzygniętych (35%), Barland uzyskał 53% poprawnych rozstrzygnięć i 12% błędnych.

Zakładaliśmy, zgodnie z doświadczeniami innych autorów⁵, że wyniki uzyskane w warunkach eksperymentalnych będą gorsze niż uzyskiwane w badaniach wykonywanych w praktyce, gdzie od wyniku badania zależeć mogą losy badanego, a zatem jego napięcie emocjonalne i motywacje są o wiele większe niż w badaniu eksperymentalnym, a zarówno pytania krytyczne, jak i kontrolne dotyczą tej samej sfery realnego życia. Zakładaliśmy także, że z uwagi na wspomniany większy z natury rzeczy ciężar gatunkowy pytania kontrolnego, częstszy będzie błąd poligraferów polegający na uznaniu za osobę prawdomówną osoby kłamliczej niż odwrotnie.

W badaniach poligraficznych, przeprowadzonych w ramach finansowanego przez Narodowe Centrum Nauki projektu badawczego pt. „Instrumentalne i nieinstrumentalne metody detekcji nieszczerości – problemy kryminalistyczne, etyczne i prawne”, do celów diagnostycznych wykorzystywano test porównania stref – Utah ZCT, opracowany przez Uniwersytet w Salt Lake City. Innym testem był test szczytowego napięcia (POT – z j. ang. *peak of tension*), ale miał zastosowanie wyłącznie demonstracyjne.

6.1. Geneza techniki Utah

Test Utah ZCT jest elementem szerszego podejścia, kompletnej techniki badawczej Utah z rodziny technik pytań porównawczych (CQT – z j. ang. *comparison questions techniques*), obejmującej nie tylko konkretny format testu (rodzaje i sekwencję pytań), ale także reguły prowadzenia wywiadu przedtestowego, omawiania i prezentowania bodźców testowych oraz metodę analizy danych testowych wraz z kryteriami decyzyjnymi.

² Por. J. Widacki: *Spór o technikę badań poligraficznych*, „Studia Prawnicze. Rozprawy i Materiały” 2016, 2 (19), s. 9–23, a także: M. Tarabuła, M. Widacki, *The Amount of Information Remembered by Perpetrator In the Context of the Application of the Guilty Knowledge Technique In Criminal Investigation – a Pilot Study*, „European Polygraph” 2016, 10, 2 (36), s. 63–76.

³ Por. J. Widacki, *Wprowadzenie do problematyki badań poligraficznych*, Warszawa 1981, s. 128.

⁴ G. Barland, *An experimental study of field techniques in lie-detection*, (niepubl.) University of Utah, 1972, obszernie opisane w: J. Widacki, *Wprowadzenie...*, op. cit., s. 127–128.

⁵ Orlansky, zestawiając wyniki badań eksperymentalnych wielu autorów, podaje, że posługując się różnymi technikami, uzyskiwali oni od 75–100% trafnych diagnoz (J. Orlansky: *An assessment of lie-detection capability*, Institute for Definitive Analysis, Research and Engineering Support Division, „Technical Report” 62–66, July 1964, cyt. za: J. Widacki, *Wprowadzenie...*, op. cit., s. 128); Abrams podaje, że w badaniach eksperymentalnych osiąga się zwykle nieco ponad 80% poprawnych diagnoz (S. Abrams, *Polygraph validity and reliability: A review*, „Journal of Forensic Sciences” 1973, 18, 4, s. 318).

Prace nad udoskonaleniem technik pytań porównawczych (kontrolnych), które zaowocowały zaprojektowaniem systemowego podejścia „Utah”, rozpoczęły się w latach 70. XX w. na Uniwersytecie w Salt Lake City (w amerykańskim stanie Utah). Zespołowi naukowców przewodniczył prof. David Raskin. Jego współpracownikami byli m.in. dr John Kircher i dr Charles Honts, a podstawę prac stanowił dotychczasowy dorobek Johna Reida, Cleve’a Backstera i Armii Stanów Zjednoczonych.

6.2. Główne założenia i struktura testu Utah ZCT

Podczas badań w ramach grantu wykorzystywano wersję testu diagnostycznego Utah ZCT o charakterze wieloaspektowym (zawierającego dwa pytania krytyczne dotyczące głównego aspektu zainscenizowanego zdarzenia – tj. oddania strzału na akademickiej strzelnicy sportowej, oraz jedno pytanie krytyczne – tzw. łączące dowody, np. o trzymanie w ręku tej broni, z której oddano strzał, czy obecność przy tej czynności). Sekwencja pytań tego testu była następująca:

I – pytanie wprowadzające – *Czy rozumie Pan/Pani, że w tym teście zadam tylko te pytania, które omówiliśmy?*

ScR – pytanie relewantne poświęcone, np. *Jeżeli to Pan/Pani oddał/a strzał na strzelnicy akademickiej w miniony czwartek – czy zamierza Pan/Pani odpowiadać szczerze na każde pytanie w tej kwestii?*

N1 – pytanie neutralne, np. *Czy jesteście w Krakowie?*

C1 – pytanie porównawcze (kontrolne)⁶, np. *Czy w przeszłości kłamał/a Pan/Pani, aby osiągnąć jakąś korzyść?*

R1 – pytanie relewantne (krytyczne), np. *Czy to Pan/Pani oddał/a ten strzał?*

N2 – pytanie neutralne

C2 – pytanie porównawcze (kontrolne)

R2 – pytanie relewantne (krytyczne), np. *Czy to Pan/Pani oddał/a strzał na strzelnicy w miniony czwartek?*

N3 – pytanie neutralne

C3 – pytanie porównawcze (kontrolne)

R3 – pytanie relewantne (krytyczne), np. *Czy trzymał/a Pan/Pani w ręku tę broń, z której oddano strzał w miniony czwartek?*

Poza zastosowaniem testu Utah ZCT jako wieloaspektowego, jest on wykorzystywany przede wszystkim jako test jednoproblemowy. Wówczas wszystkie pytania krytyczne muszą być względem siebie zależne pod względem logicznym i nie ma możliwości, by badany – kłamiąc na jedno pytanie – był równocześnie szczerzy przy drugim. Odwrotnością testów jednoproblemowych są testy wieloproblemowe (typu MGQT, DLST), gdzie pytania krytyczne dotyczą różnych zdarzeń i czynów, mają charakter przesiewowy i siłą rzeczy – nieco mniejszą dokładność niż testy diagnostyczne.

⁶ Dopuszcza się wykorzystanie pytań kontrolnych zarówno typu PLC (z j. ang. *probable lie comparison*), jak i DLC (z j. ang. *directed lie comparison*). Pytanie PLC jest tak zaprezentowane i w takim kontekście, by badany prawdopodobnie nie udzielił do końca szczerzej odpowiedzi (zazwyczaj przeczącej). Ewentualnie zakres pytania jest na tyle szeroki, by badany nie miał całkowitej pewności, że niczego nie pominął, a jeśli nawet odpowie szczerze – takie pytanie będzie wymagało większego kognitywnego zaangażowania niż przy pytaniach relewantnych (np. Czy kiedykolwiek postąpił Pan nieuczciwie?). Z kolei pytanie DLC polega na tym, że badany otrzymuje instrukcję udzielania odpowiedzi fałszywych (niezgodnych ze swoją wiedzą), co prowadzi do wywołania dysonansu poznawczego.

Przy projektowaniu testu Utah ZCT wzięto pod uwagę dwa podstawowe fakty:

- 1) Osoby wprowadzające w błąd zwykle reagują w sposób bardziej znaczący, udzielając odpowiedzi na pytania krytyczne niż na pytania kontrolne. Natomiast u osób prawdziwych występuje odwrotna sytuacja – dla nich istotniejszymi bodźcami będą pytania kontrolne;
- 2) Różnice w wielkości reakcji między pytaniami krytycznymi a kontrolnymi są w przypadku osób prawdziwych znacznie mniejsze niż w przypadku badanych nieszczerých.

Potrzebne były więc dodatkowe zabiegi, aby poprawić trafność decyzyjną i zmniejszyć liczbę potencjalnych wyników nierozstrzygniętych u osób odpowiadających na pytania krytyczne zgodnie z własną wiedzą i przekonaniem. Pytanie krytyczne w teście za każdym razem poprzedzone zostało pytaniem neutralnym i kontrolnym. Pytania neutralne wywołują niewielkie zmiany reakcji fizjologicznych. W konsekwencji – przed kolejnym pytaniem reakcje mają możliwość powrotu do linii bazowej, więc nie ma ryzyka, że reakcja na pytanie kontrolne będzie zmniejszona z powodu wysokiego pułapu linii bazowej po poprzednim bodźcu. Dopiero po pytaniu kontrolnym wprowadzane jest pytanie krytyczne. Zjawisko habituacji (stopniowego osławania z bodźcami) w ramach serii testu sprzyja „faworyzowaniu” pytań kontrolnych i łagodzi reakcje na pytanie krytyczne. Dodatkowo, jeśli reakcja na pytanie kontrolne była silna i nie powróciła do poziomu bazowego przed zadaniem pytania krytycznego – to również, zgodnie z prawem wartości wstępnych (z j. ang. *law of initial values*), osławieniu ulegnie reakcja na pytanie krytyczne.

W związku ze zjawiskiem habituacji, które występuje nie tylko w ramach danej serii, ale również między poszczególnymi seriami testu, pytania podlegają rotacji – tak, aby każde pytanie krytyczne i kontrolne pojawiło się w podobnym stopniu w różnych miejscach w sekwencji, a badany nie mógł przewidzieć, o co w danym momencie zostanie zapytany⁷. Musi znać treść wszystkich pytań przed testem, ale bez uprzedzania o ich kolejności (wyjątkiem w tym aspekcie jest jedynie test szczytowego napięcia – POT). W ten sposób każde pytanie krytyczne przynajmniej raz zostanie umiejscowione w sąsiedztwie każdego pytania kontrolnego i neutralnego. Na swoich stałych pozycjach pozostają jedynie pytania: wprowadzające (z j. ang. *introductory*) i relewantne poświęcone (z j. ang. *sacrificed relevant*). To pierwsze ma za zadanie upewnić badanego, że pytania testowe będą dotyczyły wyłącznie tych kwestii, które są przedmiotem badania, a nie jakichś innych potencjalnych problemów wywołujących obawy. Drugie natomiast pozwala oswoić badanego z obecnością pytania brzmiącego oskarżająco, a reakcja na to pytanie nie podlega późniejszej ewaluacji. Rotacje pytań w poszczególnych seriach testu przedstawiają się następująco:

- I seria: I, ScR, N1, **C1**, **R1**, N2, **C2**, **R2**, N3, **C3**, **R3**.
- II seria: I, ScR, N2, **C3**, **R2**, N3, **C1**, **R3**, N1, **C2**, **R1**.
- III seria: I, ScR, N3, **C2**, **R3**, N1, **C3**, **R1**, N2, **C1**, **R2**.

W ewentualnych kolejnych dwóch seriach – albo powtarzamy sekwencje z serii pierwszej i drugiej, albo dowolnie rotujemy pytania.

6.3. Metodyka przeprowadzania testu Utah ZCT

Badanie rozpoczynamy od wywiadu przedtestowego, który obejmuje m.in. przedstawienie celów badania, uzyskanie zgody na poddanie testom z wykorzystaniem poligrafu oraz omówienie kwestii związanych z biografią i stanem zdrowia badanego. Ekspert stwierdza,

⁷ Warto również zauważyć, że po przerwie między seriami testu wielu badanych odzwyczaja się od dotychczasowych bodźców i reakcje ponownie przybierają na sile.

czy nie istnieją ewentualne przeciwwskazania do kontynuowania badania, i informuje badanego o rejestrowaniu przebiegu czynności w systemie audiowizualnym, wraz z odpowiednią argumentacją (zasada dokumentowania, procedury kontroli jakości, sporządzenie wiernego sprawozdania). Dalej nawiązuje odpowiedni kontakt z badanym, zwracając jednocześnie uwagę na jego zdolności intelektualne i językowe. Na tym etapie w rozmowie zadaje się pytania otwarte.

Bardzo ważnym elementem wywiadu jest faza swobodnej wypowiedzi badanego (wolna narracja) na temat swojej wiedzy i związku ze sprawą będącą przedmiotem badania. Chodzi o to, by uzyskać jak najwięcej informacji bez przerywania, konfrontowania z innymi danymi i wprowadzania niepotrzebnego stresu. Badający upewnia się następnie, że badany dostrzega różnicę między faktami a podejrzeniami. W razie potrzeby pogłębia niektóre kwestie poruszone w jego wypowiedzi, zwraca uwagę na niespójności, ale robi to w dalszym ciągu w sposób niekonfrontacyjny.

Po wyjaśnieniu wszelkich wątpliwości, które mogłyby wzmacniać niepewność badanego, prosi się go o zajęcie miejsca na fotelu do badań i zakłada się czujniki. W tym czasie krótko omawia się przeznaczenie poszczególnych czujników i generalne zasady funkcjonowania autonomicznego układu nerwowego.

Kolejnym krokiem jest przeprowadzenie testu demonstracyjnego (zapoznawczego, stymulacyjnego). Jest to test szczytowego napięcia – najczęściej z wykorzystaniem liczb czy imion. Badany zapoznaje się z tym, jak wygląda procedura, oswaja się z sytuacją na tyle, żeby nie koncentrował się na sprzęcie, lecz na dalszych pytaniach. Ekspert z kolei upewnia się, że poligraf działa prawidłowo, dostraja czułości czujników i obserwuje reakcje badanego – zwłaszcza przy pytaniu, przy którym uzgodniono udzielenie odpowiedzi fałszywej. Po teście przekazuje się badanemu ewentualne uwagi (np. co do sposobu oddychania i potrzeby utrzymania stabilnej pozycji ciała), a przede wszystkim jasny komunikat, że poligraf zarejestrował typowe zmiany reakcji fizjologicznych – takie, które występują zwykle przy wprowadzaniu w błąd, gdy badany odpowiadał niezgodnie ze stanem faktycznym – i że reagował tak jak powinien przy odpowiedziach zgodnych z prawdą.

Następnie przystępuje się do omawiania pytań testowych według ustalonej kolejności – najpierw pytanie relewantne poświęcone, potem pozostałe pytania relewantne, kontrolne, neutralne i na końcu – wprowadzające. W momencie, w którym wszystko jest zrozumiałe i badany jest w stanie udzielać na każde pytanie krótkiej, jednoznacznej odpowiedzi w formie „tak” lub „nie”, rozpoczyna się etap rejestrowania danych przez poligraf. Między seriami testu wskazane jest odniesienie się w rozmowie do wszystkich pytań relewantnych i porównawczych, pamiętając o równomiernym rozłożeniu akcentów – w szczególności, aby nie skierować przesadnej uwagi na zagadnienia krytyczne. Chodzi o to, aby upewnić się, że pytania krytyczne pozostają jasne dla badanego, a pytania porównawcze – postrzegane w dalszym ciągu jako istotne, nie są bagatelizowane. Stymulacja między wykresami zmniejsza liczbę wyników fałszywych (zwłaszcza negatywnych) i nierozstrzygniętych⁸.

Po zakończeniu testu i wstępnej analizie danych badanemu umożliwia się złożenie ewentualnych dodatkowych oświadczeń w kwestiach mogących budzić wątpliwości po obu stronach. Na tym etapie możliwe jest już konfrontowanie z innymi posiadanymi informacjami, lecz nadal z zachowaniem warunków udzielania przez badanego swobodnych wypowiedzi – tak, aby ekspertyza nie przerodziła się w kodeksowo niedopuszczalne przesłuchanie.

⁸ C.R. Honts, *The Discussion of Comparison Questions Between List Repetitions (Charts) is Associated with Increased Test Accuracy*, „Polygraph”, 28(2), s. 117–123.

6.4. Ewaluacja danych testowych

Zanim powstał system analizy danych testowych Utah, korzystano z innych, niezbyt doskonałych metod, które w poważnej części opierały się na przesłankach diagnozowania i regułach decyzyjnych o wątpliwych podstawach naukowych. Dominujący w przeszłości system oceniania Backstera był bardzo skomplikowany i stosunkowo niekorzystny dla osób prawdomównych. W wyniku prowadzonych badań naukowych liczbę diagnostycznych cech graficznych zapisów reakcji na poligramach ograniczono do dziewięciu i opracowano system numerycznej ewaluacji.

Generalnie, jeśli bardziej istotne zmiany reakcji występują w danym parametrze (kanał pomiaru) przy pytaniu krytycznym, ocenę liczbową poprzedzamy znakiem minus. Natomiast przy bardziej znaczących reakcjach na porównywane pytanie kontrolne ocena jest dodatnia. W systemie Utah korzystamy z 7-pozycyjnej skali punktacji $<-3,+3>$. I tak:

- a) ocenę „0” przypisujemy, gdy zmiany reakcji są zbliżone lub nie ma takich zapisów, które można by ze sobą porównać;
- b) $+1/-1$ – gdy występuje zauważalna różnica;
- c) $+2/-2$ – silna, wyraźna różnica;
- d) $+3/-3$ – gdy występuje dramatyczna różnica, jednoznaczny zapis krzywej i jednocześnie mamy do czynienia z najsilniejszą reakcją na wykresie.

Odruch orientacyjny jest zbiorem reakcji fizjologicznych, które kierunkują organizm człowieka na zmiany w otoczeniu, które mogłyby zagrozić jego istnieniu. Wśród bodźców, które mogą wywołać taki odruch, są np. przedmioty spadające z wysokości w naszym kierunku, dziwne hałasy czy niektóre pytania zadawane w czasie badania poligraficznego. Odruch orientacyjny przejawia się czasem w aktywności współczulnej części autonomicznego układu nerwowego (pobudzającej wystąpienie jakiejś reakcji), a w innych przypadkach – w części przywspółczulnej (zazwyczaj hamującej reakcje). Na przykład: zmniejsza się ruchliwość żołądkowo-jelitowa, tempo oddychania, wydzielanie przez błonę śluzową nosa i gruczoły ślinowe; zwężają się tętniczki skórne, źrenice rozszerzają się, a gruczoły potowe zwiększają swoją aktywność.

Na podstawie badań empirycznych wyodrębniono następujące przesłanki diagnozowania dla poszczególnych kanałów zapisu:

- a) pneumo
 - czas rozpoczęcia reakcji: od wprowadzenia bodźca do 5 sekund po odpowiedzi;
 - zakres reakcji: > 3 cykle, nawet do 20 sekund, jeśli reakcja rozpoczęła się we właściwym czasie;
 - zmniejszenie amplitudy;
 - podniesienie linii bazowej;
 - bezdech;
 - spowolnienie częstotliwości.
 - przypisywanie ocen numerycznych odbywa się na wyżej opisanych zasadach ogólnych;
 - w przypadku stwierdzenia dwóch równorzędnych cech diagnostycznych zwraca się uwagę na czas trwania reakcji (odcinek krzywej do porównania musi się znajdować w przedziale od wprowadzenia bodźca do 10 kolejnych sekund).
- b) EDA / GSR
 - czas rozpoczęcia reakcji: 0,5 sekundy od wprowadzenia bodźca do 5 sekund po odpowiedzi;
 - zakres reakcji: od rozpoczęcia reakcji do powrotu do linii bazowej;

- amplituda (cecha główna mierzona od linii bazowej do szczytu reakcji);
 - pomocniczo: czas trwania i złożoność;
 - 1 = podwójna różnica w amplitudzie, ewentualnie stosunek 1,5:1 + czas trwania lub złożoność;
 - 2 = potrójna różnica w amplitudzie, ewentualnie stosunek 2,5:1 + czas trwania reakcji lub jej złożoność;
 - 3 = poczwórna różnica w amplitudzie, jednoznaczny zapis krzywej – jest to najsilniejsza reakcja na wykresie.
- c) cardio
- czas rozpoczęcia reakcji: od wprowadzenia bodźca do 5 sekund po odpowiedzi;
 - zakres reakcji: od rozpoczęcia reakcji do powrotu do linii bazowej;
 - podniesienie linii bazowej (wzrost krzywej – wyraźniejszy po stronie rozkurczowej);
 - czas trwania reakcji;
 - 1 = stosunek wielkości reakcji 1,5:1;
 - 2 = stosunek 2:1;
 - 3 = stosunek 3:1 – jest to najsilniejsza reakcja na wykresie.
- d) PLE / PPG
- czas rozpoczęcia reakcji: 2 sekundy od wprowadzenia bodźca do 5 sekund po odpowiedzi;
 - zakres reakcji: do 20 sekund;
 - redukcja amplitudy (zwężenie krzywej) i czas trwania tej redukcji;
 - oceny 1 lub 2, które można przypisać także, gdy nie ma różnicy w wielkości redukcji amplitudy, ale istnieje wyraźna różnica w czasie trwania reakcji.

Zarówno w literaturze anglojęzycznej, jak i polskiej, możemy spotkać przykładowe zbiory z diagnostycznymi wzorami reakcji⁹.

Po przypisaniu ocen liczbowych zapisom reakcji na pytania krytyczne stosuje się ustalone progi decyzyjne:

- a) gdy suma całkowita ≤ -6 wynik testu oznaczamy jako **DI** (z j. ang. *deception indicated* – **stwierdzono wprowadzanie w błąd**);
- b) gdy suma całkowita $\geq +6$ wynik stwierdzamy jako **NDI** (z j. ang. *no deception indicated* – **nie stwierdzono wprowadzania w błąd**);
- c) w przypadku pozostałych wyników uznajemy test za nierozstrzygnięty – **INC** (z j. ang. *inconclusive*).

Do ewaluacji zapisów reakcji badanego można w przypadku testu Utah ZCT z powodzeniem zastosować także inne potwierdzone naukowo manualne systemy numerycznej analizy danych testowych, np. Empiryczny System Oceniania (ESS)¹⁰. Warto dodać, że obok systemów manualnych, biegli mogą pomocniczo posilkować się coraz dokładniejszymi algorytmami komputerowymi (np. w ramach programu OSS-3).

⁹ Zob. m.in.: M. Handler, R. Nelson, *Utah Approach to Comparison Question Polygraph Testing*, „Polygraph” 2009, Vol. 38, No. 1, s. 15–30; *Współczesne standardy badań poligraficznych*, red. M. Gołaszewski, Warszawa 2013, s. 29–34.

¹⁰ Zob. m.in.: D. Słapczyńska, *Ocena zapisów testów porównań strefowych UTAH przy użyciu Empirycznego Systemu Oceny (ESS)*, „Problemy Kryminalistyki” 2013, nr 282 (4); M. Gołaszewski, *Potwierdzone naukowo systemy analizy danych testowych*, [w:] *Współczesne standardy...*, op. cit., rozdz. 3, s. 26–41.

6.5. Wartość diagnostyczna i dowodowa

Cała koncepcja badań poligraficznych techniką Utah zyskała potwierdzenie w wielu – zarówno laboratoryjnych, jak i terenowych – badaniach naukowych w ciągu ostatnich czterdziestu lat. Była przedmiotem licznych publikacji, poddawanych niezależnym recenzjom. Średnia dokładność testu diagnostycznego Utah ZCT wynosi aż 90,2–93%, zaś średni odsetek wyników nierozstrzygniętych przedstawia się na poziomie 7,3–10,7% (w zależności od rodzaju zastosowanych pytań porównawczych i systemu analizy danych testowych). Dlatego w badaniach kryminalistycznych praktykujący biegli najczęściej sięgają po ten właśnie test. W specjalnym raporcie American Polygraph Association umieszczono go na liście testów rekomendowanych do badań dowodowych¹¹. Spełnił kryteria w postaci przynajmniej dwóch badań empirycznych (oryginalnych i powtórzonych) opublikowanych w niezależnie recenzowanych czasopismach naukowych, a także kryteria dotyczące średniej dokładności (minimum 90%) i odsetek wyników nierozstrzygniętych (maximum 20%). Tożsame rekomendacje wydało również Polskie Towarzystwo Badań Poligraficznych.

Przez część środowiska akademickiego w Polsce test Utah ZCT, w połączeniu z numeryczną metodą ewaluacji zapisów na poligramach, został uznany wręcz za jedyny test przedstawiający istotną wartość dowodową¹². Choć pogląd ten wydaje się nieco przesadzony – deprecjonuje kilka innych cennych technik badawczych – to słusznie uwypukla dorobek naukowców z Uniwersytetu w Salt Lake City.

7. Wyniki badań poligraficznych

W badanej grupie 39 osób poligraferzy 11 osób zaliczyli do grupy „kłamających” (DI), 26 do grupy „niekłamających”, a wynik badania 2 osób uznali za nierozstrzygnięty (INC).

Tabela 2. Wskazania poligraferów

Kłamący (DI)	Niekłamący (NDI)	Wynik nierozstrzygnięty (INC)	Razem
11	26	2	39
28%	67%	5%	100%

Wskazując 11 osób jako „kłamające” (DI), poligraferzy 8 osób wskazali trafnie, a 3 osoby „niekłamające” (NDI) błędnie uznali za „kłamające” (DI).

Wskazując 26 osób jako „niekłamające” (NDI), w 19 przypadkach wskazali trafnie, a w 7 przypadkach błędnie osoby „niekłamające” (NDI) wskazali jako „kłamające” (DI).

W dwóch przypadkach poligraferzy nie byli w stanie jednoznacznie określić, do której grupy „kłamających” czy „niekłamających”, zaliczyć badanych (INC).

¹¹ *Meta-Analytic Survey of Criterion Accuracy of Validated Techniques*, „Polygraph” 2011, Vol. 40, No. 4. Omówienie raportu: M. Gołaszewski, *Validated Techniques and Scoring Models for PDD Test Data Analysis – Conclusions from the 2011 APA Report*, „European Polygraph” 2012, nr 4 (22).

¹² P. Herbowski, *II Międzynarodowe Seminarium Poligraficzne „Modyfikacja i standaryzacja metod badawczych stosowanych w procesowych i kadrowych badaniach psychofizjologicznych”*, „Problemy Kryminalistyki” 2010, nr 267.

Tabela 3. Zestawienie wyników eksperymentu

Wskazania	Trafne	Błędne	Nierozstrzygnięte	Razem
DI	8	3	-	11
NDI	19	6	-	25
INC	-	-	3	3
Razem	27	10	2	39

Tabela 4. Miary trafności ocen poligraficznych (czułość, swoistość, PPV, NPV oraz trafność). *N* oznacza liczebność

Metoda	N wskazań (sprawca/ niewinny)	N badanych (sprawca/ niewinny)	Czułość (N poprawnych wskazań sprawców)	Swoistość (N poprawnych wskazań niewinnych)	PPV	NPV	Trafność (N poprawnych wskazań ogółem)
Ocena z wyłączeniem wyników niekonkluzywnych	36 (11/25)	36 (14/22)	57% (8)	86% (19)	73%	76%	75% (27)
Ocena z włączeniem wyników niekonkluzywnych	39 (11/25)	39 (15/24)	53% (8)	79% (19)	62%	73%	69% (27)

Jak widać z powyższej tabeli, dokładność (trafność) poligrafu w tym badaniu wynosi 75%. Czulość oznacza, ilu spośród wszystkich sprawców zostało wykrytych. Wynosi ona 53%, co jest dość niskim wynikiem (oznacza to, że wszystkich „sprawców” wykryto tylko 53%).

Z kolei swoistość oznacza, ile spośród wszystkich niewinnych osób trafnie wskazało na-rzędzie i w prezentowanym badaniu wynosi 86%. Oznacza to, że spośród wszystkich niewinnych osób badanych wykryto ich 86%.

Dużo ciekawsze są miary PPV i NPV. PPV to wartość predykcyjna dodatnia – współczynnik pokazujący, jaki procent osób określonych jako „sprawcy” faktycznie było sprawcami. W prezentowanym badaniu PPV wynosi 73%. Oznacza to, jest 73% szansy, że dana osoba, uznana w badaniu poligraficznym za sprawcę, faktycznie jest sprawcą. Z kolei NPV to wartość predykcyjna ujemna – pozwala nam określić, ile procent osób określonych jako „niewinne” faktycznie było niewinnych. Wynosi ona 76% co oznacza, że jeżeli w badaniu poligraficznym wskazana osoba zostanie określona jako „niewinna”, to jest 76% szansy na to, że faktyczni jest niewinna.

Jeśli uwzględnić sytuacje, kiedy poligraferzy wstrzymali się od wskazania (INC, przypadki nieoznaczone), wówczas wszystkie współczynniki spadają.

Marcin Tarabuća, Michał Widacki

The Amount of Information Remembered by the Perpetrator in the Context of the Application of the Guilty Knowledge Technique in Criminal Investigation – a Pilot Study*

Количество информации, сохраненной в памяти правонарушителя в контексте использования тестов ГКТ в ходе расследования уголовного дела: экспериментальное исследование

Key words: Guilty Knowledge Technique, GKT or CQT

Despite the fact that the Guilty Knowledge Technique [Lykken 1959, Lykken 1960], or GKT, originated more than five decades ago, its validity is still debatable, especially when compared to other polygraph techniques.

Partisans of GKT superiority to other techniques, especially to Control Question Technique, support their opinion on the high percentage of correct results (up to 100% in some studies) coupled with a relatively low count of inconclusive indications, or even their lack [Lykken 1973, Elaad et al. 1992, Elaad 1998].

Members of this group believe that the GKT technique provides much more protection for innocent subjects, because, unlike the CQT, the polygrapher does not ask directly about perpetration of a crime during the procedure, but instead he verifies the subject's knowledge about all distinctive aspects of the case – in this way reducing the chances of a random reaction to critical question, which could be interpreted as a deliberate lie made by truly innocent subjects [Krapohl et al. 2009]. The Guilty Knowledge Technique is also believed to provide more solid methodological background than the CQT [Lykken, 1974; Ben-Shakkar & Elaad 2002].

On the other hand followers of the CQT technique claim that from diagnostic point of view it cannot match the latest forms of the Control Questions technique [APA Meta-Analytic Survey 2011, Gołaszewski 2012, Widacki 2014]. Superiority of the CQT may also lay in the broader spectrum of its potential application [Elaad 1990, Podlesny 1994, Podlesny 2003] – its effectiveness does not rely on the existence of multiple distinctive details of the case known only to the investigators. Some problems with the distinction between perpetrators and witnesses (who have some knowledge about the case as well) have also been indicated [Konieczny et al. 1984, Bradley & Warfield 1986].

* M. Tarabuća, M. Widacki, *The Amount of Information Remembered by the Perpetrator in the Context of the Application of the Guilty Knowledge Technique in Criminal Investigation – a Pilot Study*, "European Polygraph" 2016, Vol. 10, No. 2(36), s. 111–116, DOI 10.1515/ep-2016-0008. Project DEC-2013/11/B/HS5/03856 funded by National Science Centre.

Followers of the CQT also argue about the theoretical base of the Guilty Knowledge Technique, especially about the assumption that the perpetrator is in a state of high consciousness during the act, and because of that has the ability to remember fully the whole event with high amount of details. An argument has been made that every single perpetrator of a crime is more or less stressed during the critical moments of the event. The presence of stress during a crime may reduce the level of offender's perception [Christianson 2007] and result in a possibility that perpetrators do not remember many details of the crimes that – from the perspective of the theoretical background to the GKT – they are expected to remember [Widacki 2011]. This argument is particularly interesting because, if accurate, it can discredit the application of the Guilty Knowledge Technique in criminal investigation, and consequently also its very right of existence.

In the light of the above, before any comparison of validity between GK and CQ techniques can be made, it is necessary to determine in a staged event whether subjects are able to remember properly a sufficient amount of details for the Guilty Knowledge Technique to be used effectively.

Method

Forty (40) subjects (students of Andrzej Frycz Modrzewski Kraków University, aged from 21 to 27) were divided in two equal groups: A ("perpetrators") and B ("witnesses"). Members of both groups were arranged into 20 "perpetrator –witness" pairs, and all of them duly participated in an activity prepared for the needs of the experiment. After receiving their instructions, each pair have entered a darkened shooting range where the "perpetrator" had 7 seconds to assume his or her place in the shooting range and take the blank gun. After that time, a light beam was activated and illuminated the rotating shooting target with the picture below placed on it, 4m away from the shooting range. The photo (80 × 60 cm) featured the "victim": a young woman standing in quite a dark room and talking on a mobile phone.



Photo 1. The picture used in the experiment.

From that moment, the "perpetrator" had 10 seconds to make one shot from the blank gun at the target, aiming to "kill" the "victim". After the time, the target began to rotate automatically to prevent further exposition. The "witness", unaware of the instructions given to the perpetrator, had to observe passively the whole event. After the target began its rotation, subjects were asked to leave the room and separately asked to fill in a questionnaire, where they first determined the level of stress generated during the experiment and then described shortly the whole event from their point of view. This was

followed by answering 11 questions. The author of the questionnaire believes that they indicated the most distinctive details of the picture. They related to:

- the gender and age of the “victim”
- situation, in which the “victim” was “caught”
- characteristic background details of the picture
- “victim’s” hair color;
- “victim’s” cloths and other details
- the objects in the “victim’s” hands
- two particular, highlighted background elements in the pictured room (a wooden bookcase to the right from the “victim”, and candlesticks with candles on the wall on the left).

The questions were to determine the amount of information that the subjects remembered while being exposed to the picture, and would be considered a starting point to develop polygraph tests using the GKT technique. The dramatic scenario of the experiment (unknown to the last moment, with little time to prepare and shoot blank gun, and also the loud noise accompanying the shooting) was developed to generate a relatively high level of stress, especially in the “perpetrators”.

Results

In the questionnaires filled after the experiment all subjects described the course of the event without much detail but correctly. Descriptions of the picture placed on the shooting target were less accurate. Reasons for that are different, and they will be presented later in this article. The stress level generated by the event as declared by subjects (on a scale 1-10, where 1 is totally free of stress and 10 fully stressed) was distributed as shown in the table below:

Table 1. Distribution of declared levels of stress in both groups.

Group	Level of stress declared by a subject		
	1–3 (low stress)	4–6 (medium stress)	7–10 (high stress)
A (“perpetrators”)	9 subjects (45%)	5 subjects (25%)	6 subjects (30%)
B (“witnesses”)	11 subjects (55%)	5 subjects (25%)	4 subjects (15%)

At the first sight, the values seem to be very similar in both groups. The chi-square (χ^2) test value in this case is 0.6 and lies outside the acceptance region for a significance level of 0.05, in the context of the critical value of chi-square distribution with two degrees of freedom – 5.991. With respect to the above, the null hypothesis cannot be rejected, which means that the amount of stress generated by the experiment cannot be considered distinctive for members of the two experimental groups.

Answers to the eleven questions about the distinctive elements of the picture used in the experiment allowed to determine the amount of information effectively remembered by subjects participating in the event.

Table 2. Distribution of the number of details (information) remembered in both groups.

Group	Number of well-remembered details:		
	0–3	4–7	8–11
A (“perpetrators”)	8 individuals (40%)	11 individuals (55%)	1 individual (5%)
B (“witnesses”)	2 individuals (10%)	12 individuals (60%)	6 individuals (30%)

The average number of remembered details of the picture exposed during the experiment is 3.8 in group A (“perpetrators”) and 6.4 in group B (“witnesses”). The chi-square test value is 7.27 and lies in the acceptance region for the significance level of 0.05, because the critical value of chi-square distribution with two degrees of freedom is 5.991. With respect to the above, there are grounds to reject the null hypothesis in this case and the distinction between the two groups of subjects based on the number of details remembered is statistically relevant. The role in the experiment affected the ability of remembering details well, independently from the subject’s declared level of stress.

Due to the large difference between the declared levels of stress (the lowest recorded value being 1 and the highest – 8) it seems reasonable to compare values of stress with the number of details remembered by the subjects regardless of their role in the experiment. The comparison of all 40 subjects participating in the experiment is presented below:

Table 3. Distribution of the number of details remembered broken by the declared level of stress in members of Group A and B together.

Declared level of stress	Number of details remembered by individuals		
	0–3	4–7	8–11
1–3 (low stress)	3 individuals (7.5%)	10 individuals (25%)	7 individuals (17.5%)
4–6 (medium stress)	2 individuals (5%)	8 individuals (20%)	-
7–10 (high stress)	5 subjects (12.5%)	5 subjects (12.5%)	-

The average amount of details remembered by the subjects who declared low stress level was 6.25, medium stress level allowed to obtain on average 4.3 details, and high level of stress – only 3.6 of details in the exposed picture. The chi-square test value for these results is 11.574 and the critical value of chi-square distribution with four degrees of freedom is 9.488. The resulting value therefore lies within the acceptance region for the level of 0.05, and the null hypothesis can be rejected. Therefore, with the 0.5 level of significance, it can be stated that there is a statistically relevant relationship between the subject’s level of stress and the amount of remembered details of the event, regardless of affiliation to group A or B.

Pilot polygraph examination

A decision was reached to run a pilot project using a group of four subjects to test the conditions (both rooms and equipment) required for running the examinations. The group consisted of people participating in the experiment described above. The subjects included two from the group of the “witnesses”, one person from the group of the “perpetrators”, and one who was not connected to the event. The polygrapher was given the task to use polygraph examinations to determine who belonged to which group.

The examination made use of CQT tests, as proper use of GKT tests was impossible for a number of reasons. First, the experiment took place more than six months before the planned examination, and the knowledge of the event became destroyed in participants in the experiment, and the differences in the way the event was remembered between the witnesses and the perpetrator was possible. Consequently, which is another argument, the knowledge of the perpetrators and witnesses of the event became levelled, the only difference between the witness and the perpetrator being the fact that the perpetrator held the gun in his hand and shot. Let a good example of portraying the blurring of the differences be the fact that neither the perpetrator nor the witness remembered what weapon was used, yet both witness and the perpetrator remembered perfectly well what the target at the shooting range was. All this resulted in the lack of sufficient characteristic differences in the features of the event between the knowledge of the witness and the perpetrator, which made it impossible to use GKT tests.

For the reason above, a CQT technique was used, to be precise the latest development in the CQT family, namely the UTAH ZCT. The test was developed in the option that contains control questions about Directed Lie Control (DLC). The examination made use of two UTAH ZCT DLC tests. The first was to check whether the subject is a witness, and the second was to test whether the examinee is the perpetrator. NDI results obtained in both tests meant that the person was not connected to the event. If the first test produced NDI and the second DI, the subject was believed to be the perpetrator. Analogously, with NDI being the result of the first and DI in the second test, the subject was believed to be a witness.

Table 4. The questions used in the polygraph examination.

Witness Test	Question Type	Perpetrator Test
Are you sure I am going to ask only the questions we have discussed?	SYMPTOMAT IC	Are you sure I am going to ask only the questions we have discussed?
Are you going to answer the questions concerning the event at the shooting range truthfully?	CRITICAL (Relevant)	Are you going to answer the questions concerning the event at the shooting range truthfully?
Are you sitting on a chair?	NEUTRAL	Are you sitting on a chair?
Have you ever lied to a person who trusted you?	CONTROL (Comparison)	Have you ever lied to a person who trusted you?
Did you witness a shot being fired at the shooting range?	CRITICAL (Relevant)	Did you witness a shot being fired at the shooting range?
Are you wearing shoes?	NEUTRAL	Are you wearing shoes?
Have you ever cheated at the exam?	CONTROL (Comparison)	Have you ever cheated at the exam?
Were you at the shooting range when the shot was fired?	CRITICAL (Relevant)	Did you have a gun in your hands on that day?
Are we at a university?	NEUTRAL	Are we at a university?
Have you ever said something derogatory about another person when they couldn't hear?	CONTROL (Comparison)	Have you ever said something derogatory about another person when they couldn't hear?
Did you see the person who fired the shot at the shooting range?	CRITICAL (Relevant)	Did you fire a shot at the shooting range on that day?

Results of the pilot study:

Polygraph results obtained were ESS (Empirical Score System) scored. For tests analysing single issue (ZCT), the system features the following decision thresholds: To classify the subject as deceptive (DI – Deception Indicated), the total test score must amount at least to -4, or any of the spots needs to reach at least -7. If the global score is +2 or greater, the person classifies as NDI (No Deception Indicated). In the remaining cases we speak of inconclusive (INC) results. The results of all the tests are presented in the table below. The table provides not only the aggregated results, but also those of spot analysis, and evaluation of individual reactions to specific questions.

Table 5. Results of individual tests.

TYPE OF TEST: WITNESS				
subject A		R1	R2	R3
	PNEUMO	0	0	0
	EDA	2	2	2
	CARDIO	0	0	0
SPOT	I	2	2	2
		R1	R2	R3
	PNEUMO	0	0	0
	EDA	2	2	2
	CARDIO	-2	0	0
SPOT	II	0	2	2
		R1	R2	R3
	PNEUMO	0	0	0
	EDA	2	-2	-2
	CARDIO	1	0	-1
SPOT	III	3	-2	-3
		R1	R2	R3
TOTAL	8	5	2	1
TEST RESULT:		NDI		

TYPE OF TEST: WITNESS				
subject B		R1	R2	R3
	PNEUMO	0	0	0
	EDA	-2	-2	-2
	CARDIO	-1	-1	-1
SPOT	I	-3	-3	-3
		R1	R2	R3
	PNEUMO	0	0	0
	EDA	2	2	-2
	CARDIO	-1	0	0
SPOT	II	-3	2	-2
		R1	R2	R3
	PNEUMO	0	0	0
	EDA	-2	2	-2
	CARDIO	-1	-1	1
SPOT	III	-3	1	-1
		R1	R2	R3
TOTAL	-15	-9	0	-6
TEST RESULT:		DI		

TYPE OF TEST: WITNESS				
subject C		R1	R2	R3

TYPE OF TEST: PERPETRATOR				
subject A		R1	R2	R3
	PNEUMO	0	0	0
	EDA	2	0	2
	CARDIO	1	1	1
SPOT	I	3	1	3
		R1	R2	R3
	PNEUMO	0	0	0
	EDA	0	0	-2
	CARDIO	0	-1	1
SPOT	II	0	-1	-1
		R1	R2	R3
	PNEUMO	0	0	0
	EDA	0	0	0
	CARDIO	0	1	0
SPOT	III	0	1	0
		R1	R2	R3
TOTAL	6	3	1	2
TEST RESULT:		NDI		

TYPE OF TEST: PERPETRATOR				
subject B		R1	R2	R3
	PNEUMO	0	0	0
	EDA	-2	-2	2
	CARDIO	0	0	1
SPOT	I	-2	-2	3
		R1	R2	R3
	PNEUMO	0	0	0
	EDA	0	0	-2
	CARDIO	0	-1	1
SPOT	II	0	-1	-1
		R1	R2	R3
	PNEUMO	0	0	0
	EDA	0	-2	-2
	CARDIO	-1	-1	1
SPOT	III	-1	-3	-1
		R1	R2	R3
TOTAL	-2	-1	-2	1
TEST RESULT:		INC		

TYPE OF TEST: PERPETRATOR				
subject C		R1	R2	R3

	PNEUMO	0	0	0
	EDA	2	-2	-2
	CARDIO	0	0	0
SPOT	I	2	-2	-2
		R1	R2	R3
	PNEUMO	0	0	0
	EDA	-2	2	-2
	CARDIO	-1	-1	1
SPOT	II	-3	1	1
		R1	R2	R3
	PNEUMO	0	0	0
	EDA	0	-2	0
	CARDIO	0	0	0
SPOT	III	0	-2	0
		R1	R2	R3
TOTAL	-5	-1	-3	-1
TEST RESULT:		DI		

	PNEUMO	0	0	0
	EDA	0	0	-2
	CARDIO	0	1	0
SPOT	I	0	1	-2
		R1	R2	R3
	PNEUMO	0	0	0
	EDA	-2	2	2
	CARDIO	-1	0	1
SPOT	II	-3	2	3
		R1	R2	R3
	PNEUMO	0	0	0
	EDA	0	2	-2
	CARDIO	1	0	-1
SPOT	III	1	2	-3
		R1	R2	R3
TOTAL	-1	-2	3	-2
TEST RESULT:		INC		

TYPE OF TEST:	WITNESS			
subject D		R1	R2	R3
	PNEUMO	0	0	0
	EDA	0	-2	2
	CARDIO	1	-1	0
SPOT	I	1	-3	2
		R1	R2	R3
	PNEUMO	0	0	0
	EDA	-2	2	2
	CARDIO	0	1	0
SPOT	II	-2	3	2
		R1	R2	R3
	PNEUMO	0	0	0
	EDA	2	-2	-2
	CARDIO	0	-1	-1
SPOT	III	2	-3	-3
		R1	R2	R3
TOTAL	-3	1	-3	-1
TEST RESULT:		INC		

TYPE OF TEST:	PERPETRATOR			
subject D		R1	R2	R3
	PNEUMO	0	0	0
	EDA	-2	0	0
	CARDIO	-1	1	-1
SPOT	I	-3	1	-1
		R1	R2	R3
	PNEUMO	0	0	0
	EDA	-2	-2	2
	CARDIO	-1	-1	-1
SPOT	II	-3	-3	1
		R1	R2	R3
	PNEUMO	0	0	0
	EDA	0	-2	-2
	CARDIO	0	0	-1
SPOT	III	0	-2	-3
		R1	R2	R3
TOTAL	-11	-6	-4	-3
TEST RESULT:		DI		

Discussion

The experiment failed to achieve the situation, in which “perpetrators” of crime could reach a significantly higher level of stress than members of the “witnesses” group. Despite that the experiment indicated the existence of clear and statistically important difference between the number of details in the picture remembered by subjects who shot at it and by ones who only observed the whole event

passively. The difference may result from factors other than stress itself. The conclusion that can be made from the descriptions made by participants in the study is that the “perpetrators” (most of whom had never fired a gun before) focused their concentration mostly on the correct completion of the task, which was to shoot the blank gun. Coupled with the very short time of exposure to the image, this circumstance did not let the “perpetrators” remember perfectly all the details of the picture, and for that reason they often only picked basic information (e.g. age or gender of the “victim”, however some “perpetrators” also found these details a problem).

“Witnesses” on the other hand, had an opportunity to concentrate more on the picture during its 10-second exposition, because they had no other activity assigned for that time.

This aside, the research showed a connection between the level of stress reached during the experiment and the remembered level of detail concerning the actions. With the results of all subjects taking part in the experiment recapitulated, it can be estimated that with the increasing level of stress, the number of correctly remembered details diminishes. Regardless of the role played in the experiment, the experienced stress and its level clearly influence the quantity of details remembered from a certain event.

Results of the experiment cannot, however, substantiate a statement that perpetrators possess more specific knowledge of details of the crime. Outcomes are rather opposite: the need to focus concentration to accomplish specific tasks may result in the perpetrator retaining less information about details of a certain event than its witness.

There is another result worth indicating: both the “perpetrators” and “witnesses” of the simulated event remembered only little information, as the average result for the two groups was 3.8 and 6.4 respectively. That level of detail remembered about the event raises doubt about the potential distinction between the “perpetrators” and “witnesses” of an event by using the GKT polygraph technique. In addition, the “witnesses” who remembered the picture much better than “perpetrators” may be qualified falsely as perpetrators of presented crime because of their better knowledge of the event.

The experiment was designed to simulate the event in which the victim and the entire surrounding are completely unfamiliar to both the perpetrator and the witness. It can therefore be presumed that if participants of the event were familiar with the victim and crime scene, the level of detail remembered would be much higher.

A relatively small group of subjects (40 people) does not allow to issue any categorical statements about the cognitive value of this experiment. It seems necessary to conduct further research in this area on a much larger scale that would allow a more reliable analysis of the investigated phenomena, and provide more reliable conclusions as result.

Further studies in the area should attempt to generate more emotional involvement of participants of the experiment to generate more consistent stress reactions. In this regard, it seems appropriate to develop a pre-study narrative, which in this experiment was limited to a brief explanation of each subject’s role in the experiment. It is also possible that changing the form of exposure of the “victim” could improve the subjects’ responses; therefore a dummy could be used for this purpose instead of a photo.

It is also necessary to reinforce the role of the perpetrators in further studies, e.g. by making them more familiar with the weapon and its elements, or asking to perform some other tasks that the “witnesses” would be unaware of. In this way, the “perpetrators” would be able to obtain certain information not available to the “witnesses”, which could be useful in determining the role of a particular individual by subjecting him or her to a polygraph examination.

Analysing the results of the pilot experiment conducted, one clearly and immediately sees that it was not easy to tell the perpetrator apart from a witness using polygraph in this experiment. On the other hand, a decision which of the subjects was not connected to the case at all was incontrovertible. This may be an argument supporting the view expressed by the authors of the amendment to the code, who refer to the polygraph as a method used to the so-called “reduction of the number of suspects”. What remains a problem is distinguishing witnesses from perpetrators in the test group. There are a number of reasons for that. The first is poor motivation of the subjects to the experiment: partici-

pants in the project did not receive any reward for “deceiving the polygraph”. The other question was the fact that the instruction for the perpetrator and witness concerning the use of the blank gun was the same. The perpetrator was instructed about the weapon in the presence of the witnesses, who for that reason spent as much time same time watching the weapon, observing also the perpetrator and remaining at the site of the experiment (shooting range), which must have had an influence on blurring of the borders between the roles of different groups of subjects.

References

- Ben-Shakhar G., Elaad E., *The Guilty Knowledge Test (GKT) as an Application of Psychophysiology: Future Prospects and Obstacles*, [in:] M. Kleiner (ed.), *Handbook of Polygraph Testing*, Academic Press 2002, 87–102.
- Bradley M.T., Warfield J.F., *Innocence, Information and the Guilty Knowledge Test in the Detection of Deception*, *Polygraph* 1986, 15, 3, 183–194.
- Christianson S.Å., Freij I., Von Vogelsang E., *Searching for Offenders' Memories of Violent Crimes* [in:] Christianson S.Å. (ed.), *Offenders' Memories of Violent Crimes*, John Wiley and Sons Ltd. 2007.
- Elaad E., *Detection of Guilty Knowledge in Real-Life Criminal Investigations*, *Journal of Applied Psychology* 1990, 75, 5, 521–529.
- Ellad E., *The Challenge of the Concealed Knowledge Polygraph Test*, *Expert Evidence* 1998, 6, 3, 161–187.
- Elaad E., Ginton A., Jungman N., *Detection Measures in Real-Life Criminal Guilty Knowledge Tests*, *Journal of Applied Psychology* 1992, 77, 5, 757–767.
- Gaines K.H., *Utility and Numerical Evaluation of the Guilty Knowledge Test*, *Polygraph* 2013, 43, 2, 173–190.
- Golaszewski M., *Validated Techniques and Scoring Models for PDD Test Data Analysis – Conclusions from the 2011 APA Report*, *European Polygraph* 2012, 6, 4, 227–240.
- Krapohl D.J., McCloughan J.B., Senter S.M., *How to Use the Concealed Information Test*, *Polygraph* 2009, 38, 1, 34–49.
- Lykken D.T., *The GSR in the Detection of Guilt*, *Journal of Applied Psychology* 1959, 43, 6, 385–388.
- Lykken D.T., *Psychology and the Lie-Detector Industry*, *American Psychologist* 1974, 29, 10, 725–739.
- Lykken D.T., *The Validity of the Guilty Knowledge Technique: The Effects of Faking*, *Journal of Applied Psychology* 1960, 44, 4, 258–262.
- Konieczny J., Fraś M., Widacki J., *Pochodzenie ukrytej informacji a niektóre cechy osobowości w badaniu poligraficznym*, *Archiwum Medycyny Sądowej i Kryminalistyki* 1984, XXXIV, 1, 25–30.
- MacLaren, V.V., *A Quantitative Review of the Guilty Knowledge Test*, *Journal of Applied Psychology* 2001, 86, 4, 674–683.
- Meta-Analytic Survey of Criterion Accuracy of Validated Polygraph Techniques. Report Prepared for the American Polygraph Association Board of Directors*, *Polygraph Special Edition* 2011, 40, 4.
- Podlesny J.A., *Is the Guilty Knowledge Polygraph Technique Applicable in Criminal Investigations? A Review of FBI Case Records*, *Polygraph* 1994, 23, 1, 85–94.
- Podlesny J.A., *A Paucity of Operable Case Facts Restricts Applicability of the Guilty Knowledge Technique in FBI Criminal Polygraph Examinations*, *Forensic Sciences Communication* 2003, 5, 3.
- Pollock Ph. H., *When the Killer Suffers: Post-Traumatic Stress Reactions Following Homicide*, *Legal and Criminological Psychology* 1999, 4, 2, 185–202.
- Widacki J.(ed.), *Badania poligraficzne w Polsce*, Oficyna Wydawnicza AFM 2014.
- Widacki J., *Logical Identity of Conclusions from Polygraph Testing Performed in Control Questions test (CQT) and Guilty Knowledge Test (GKT) Techniques*, *European Polygraph* 2011, 5, 1, 5–10.

- Widacki J., *Wprowadzenie do problematyki badań poligraficznych*, Wydawnictwo MSW, Warszawa 1981.
- Widacki J., *W sprawie wyboru techniki badania poligraficznego. Czy technika oparta na testach GKT (CIT) jest lepsza od techniki opartej na testach CQ?*, *Problemy Kryminalistyki* 2011, 273, 5–10.
- Wójcikiewicz J., *CIT czy CQT?*, *Problemy Kryminalistyki* 2012, 275, 15–18.
- Wójcikiewicz J., Kulicki M., *Ekspertyza wariograficzna*, [in:] Wójcikiewicz J. (ed.), *Ekspertyza sądowa. Zagadnienia wybrane*, 2nd ed, Oficyna Wolters Kluwer Business, Warszawa 2007.

Jan Widacki

Spór o technikę badań poligraficznych*

1. W czasie, gdy prawnicy toczyli spory o dopuszczalność badań poligraficznych w procesie karnym i o znaczenie dowodowe wyników tych badań, w środowisku poligraferów i psychologów toczyły się spory o to, która technika badań poligraficznych jest lepsza i łatwiejsza do pogodzenia z wymogami procesu karnego. Technika pytań kontrolnych (właściwie: techniki pytań kontrolnych, bo jest ich wiele), czy technika Lykkena, w Polsce nazywana „techniką wiedzy o czynie”, lub... „techniką Kulickiego” (sic!)¹. Spór ten, choć niekiedy w formie nieco karykaturalnej², toczył się także w Polsce aż do ostatnich lat³. Niektórzy uczestnicy dyskusji nie wiedzieli nawet, że dyskutują nad koncepcją Lykkena (por. niżej) przypisując ją jednemu z polskich autorów. Sprawa w sumie jest dość żenująca i dająca nienajlepsze świadectwo autorom niektórych podręczników⁴. Warto wiedzieć, że przez pewien okres poligraferzy Policji w Polsce wykonywali badania poligraficzne wyłącznie techniką Lykkena (nazywając ją „techniką wiedzy o czynie”). Sądzone nawet, że „metoda Kulickiego”, w przeciwieństwie do technik pytań kontrolnych, jest do pogodzenia z polskim kpk⁵. O rzekomej wyższości tej techniki nad technikami pytań kontrolnych zupełnie niedawno przekonywali Wójcikiewicz⁶ i Wilk⁷. O ile więc na świecie sprawa jest jednoznacznie rozstrzygnięta⁸, w Polsce ten dziwny spór jest ciągle aktualny.

2. Należy w tym miejscu przypomnieć, że od lat 20. XX wieku, kiedy po raz pierwszy wykorzystano badanie poligraficzne w śledztwie, aż do roku 1947, kiedy John Reid zaproponował nową technikę (por. niżej), stosowano wyłącznie technikę zwaną dziś „klasyczną” (albo „techniką Keelera”) ⁹. Technika ta posługiwała się testami złożonymi z kilkunastu pytań, z których mniej więcej połowa była „związanych” ze sprawą, w literaturze polskiej zwanych także „krytycznymi” (ang. relevant question), oraz pytań niezwiązanych ze sprawą („obojętnych”, irrelevant question). Przy pomocy takich testów amerykańanie badali między innymi tych jeńców niemieckich, którzy deklarowali po wojnie współpracę

* J. Widacki, *Spór o technikę badań poligraficznych*, „Studia Prawnicze. Rozprawy i Materiały” 2016, nr 2 (19), s. 9–23. Publikacja powstała w ramach projektu sfinansowanego ze środków Narodowego Centrum Nauki przyznanych na podstawie decyzji numer DEC-2013/11/B/HS5/03856.

¹ B. Młodziejowski, W. Brzęk, J. Moszczyński, *Kryminalistyka*, Warszawa 2006, s. 275–276.

² Por. np. M. Kulicki, *Niepokojąca polemika*, „Gazeta Prawnicza” 16.12.1975, nr 24, s. 6, *idem*, *Kryminalistyka. Wybrane problemy teorii i praktyki śledczo-sądowej*, Toruń 1994.

³ Por. np. J. Wójcikiewicz, *CIT czy CQT*, „Problemy Kryminalistyki” 2012, nr 275; J. Widacki, *W sprawie wyboru techniki badania poligraficznego. Czy technika oparta na testach GKT (CIT) jest lepsza od techniki opartej na testach CQT?*, *Problemy Kryminalistyki* 2011, nr 273.

⁴ Sprawę szczegółowo przedstawiona została w: J. Widacki, *Badania poligraficzne w polskich podręcznikach kryminalistyki*, [w:] *Badania poligraficzne w Polsce*, red. J. Widacki, Oficyna Wydawnicza AFM, Kraków 2014.

⁵ B. Młodziejowski, W. Brzęk, J. Moszczyński, *op. cit.*, s. 276.

⁶ J. Wójcikiewicz, *op. cit.*

⁷ *Kryminalistyka. Przewodnik*, red. D. Wilk, Toruń 2013.

⁸ Wedle standardów American Polygraph Association, technika Lykkena jest dopuszczalna, jednak jest rekomendowana do stosowania w celach wykrywczych, a nie dowodowych! (Mata-Analytic Survey of Criterion Accuracy of Validated Techniques, Polygraph 2011, 40, 1).

⁹ Por. J. Widacki, *Wprowadzenie do problematyki badań poligraficznych*, Wydawnictwo MSW, Warszawa 1981, s. 56 i nast.

z siłami okupacyjnymi na terenie Niemiec. Pytania krytyczne („związane”) były m.in. takie: „Czy należałeś do partii nazistowskiej?”, „Czy służyłeś w Gestapo?”, „Czy służyłeś w SS?” itp. Pytania obojętne („niezwiązane”) były tego rodzaju: „ Czy znajdujemy się na Rhode Island?”, „Czy dziś jest czwartek?” Na pytania krytyczne (związane) badani odpowiadać mieli „nie”, na obojętne („niezwiązane”) – „tak”. W ramach jednego badania ten sam test, choć niekoniecznie z pytaniami w tej samej kolejności, powtarzano zwykle dwu lub trzykrotnie. Od końca lat 20. XX wieku testy klasyczne uzupełniano testami, które później otrzymały nazwę „testów szczytowego napięcia” (peak of tension, POT)¹⁰. Testy te skonstruowane są na innej zasadzie niż testy „klasyczne”. W „testach szczytowego napięcia” pytanie krytyczne (jedno w teście) umieszczone jest wśród innych podobnych pytań. Założenie jest takie, że nieszczerzy badany, zamierzający kłamać, wie które z pytań jest krytyczne. Dla osób nieznających szczegółów przestępstwa, wszystkie pytania są równie „krytyczne”. Tak na przykład, gdy zabójstwa dokonano za pomocą noża, a podejrzany poddany badaniom twierdzi, że nie tylko nie jest sprawcą, ale nawet nie zna żadnych szczegółów przestępstwa, o popełnienie którego jest podejrzany, można dla niego ułożyć taki test „szczytowego napięcia”:

- 1) Czy uderzyłeś tego mężczyznę siekierą?
- 2) Czy uderzyłeś tego mężczyznę łomem?
- 3) Czy strzeliłeś do tego mężczyzny?
- 4) Czy uderzyłeś tego mężczyznę nożem? (pytanie krytyczne, „klucz”)
- 5) Czy udusiłeś tego mężczyznę?
- 6) Czy uderzyłeś tego mężczyznę łyżką do opon?

Takie testy stosowali zarówno Leonard Keeler, jak i Clarence D. Lee. Ten ostatni, jak podają Krapohl i Shaw¹¹, w liście do szefa FBI przekonywał do możliwości oparcia całego badania poligraficznego wyłącznie na testach POT. Leonard Keeler, obok typowych testów POT, skonstruował nowy rodzaj tych testów, tak zwany test POT-poszukiwawczy („searching peak”), przy którym zarówno poligraf, jak i osoba niezwiązana ze sprawą nie znają prawidłowego „klucza”. Taki test stosowano (i jest stosowany nadal) do celów poszukiwawczych. Gdy na przykład nie ma pewności gdzie sprawca ukrył dowody przestępstwa, a z okoliczności sprawy wynika, że najprawdopodobniej je ukrył i badanie ma pomóc w ustaleniu miejsca ich ukrycia, taki test wygląda na przykład tak:

- 1) Czy schowałeś to u siebie w domu?
- 2) Czy schowałeś to w domu swojej dziewczyny?
- 3) Czy schowałeś to w miejscu pracy?
- 4) Czy schowałeś to w domu swoich rodziców?
- 5) Czy schowałeś to w domu u brata?
- 6) Czy schowałeś to w innym miejscu, o które jeszcze nie pytałem?

W roku 1947 John Reid, na łamach wydawanego przez Northwestern University w Chicago czasopisma „Journal of Criminal Law and Criminology”¹², ogłosił krótki artykuł: *A revised questioning technique In lie-detection tests*. W artykule tym autor zaproponował wprowadzenie do testów poligraficznych, obok pytań krytycznych i obojętnych, trzeciego rodzaju pytań: pytania kontrolne. Miały być one tak zbudowane, aby badany odpowiadając na nie powiedział „nie”, ale mówiąc „nie”, aby albo świadomie skłamał, albo nie był pewien czy nie skłamał. Ocena zapisu poligraficznego miała się teraz skupiać na porównaniu reakcji na pytanie krytyczne z reakcją na pytanie kontrolne. Wedle założeń Reida osobnik, który na pytanie krytyczne odpowiada nieszczerze, bardziej zareaguje na to pytanie krytyczne niż na pytanie kontrolne. I na odwrót. Osoba, która na pytanie krytyczne szczerze zaprzeczy, nie zareaguje na to pytanie, a zareaguje na pytanie kontrolne. Tak powstała technika pytań kontrolnych Reida, stosowana powszechnie w USA do końca lat 60. XX wieku, kiedy to pojawiły się nowsze techniki pytań kontrolnych (najpierw technika Backstera, później dalsze rozwinięcia tej techniki, w tym między innymi: Federal Zone Comparison Technique, Utah Zone Comparison Tech-

¹⁰ Por. *ibidem*, s. 62 i nast.

¹¹ D. Krapohl, P. Shaw, *Fundamentals of polygraph practice*, Elsevier, Academic Press, 2015, s. 27.

¹² „Journal of Criminal Law and Criminology” 1947, 37, 542.

nique). W Polsce technika Reida stosowana była powszechnie do końca lat 80.¹³, a przez niektórych poligraferów bywa stosowana nawet i dziś, choć jest to już technika zdecydowanie przestarzała.

Układ testu pytań kontrolnych Reida wyglądał następująco¹⁴:

- 1) Czy znajdujemy się w Chicago? (obojętne)
- 2) Czy nazywają cię „Rudy”? (obojętne)
- 3) Czy ostatniej soboty strzelałeś do Johna Jonesa? (krytyczne)
- 4) Czy ty palisz? (obojętne)
- 5) Czy zabiłeś Jonesa w nocy w ostatnią sobotę? (krytyczne)
- 6) Czy oprócz tego, co mi powiedziałeś, ukradłeś coś jeszcze? (kontrolne)
- 7) Czy ty chodziłeś do szkoły? (obojętne)
- 8) Czy w ostatnią sobotę w nocy zabrałeś Jonesowi zegarek? (krytyczne)
- 9) Czy wiesz kto strzelał do Johna Jonesa? (krytyczne)
- 10) Czy kiedykolwiek ukradłeś coś w miejscu pracy? (kontrolne)

W ramach techniki Reida zalecane było, obok powtarzanych dwu lub trzykrotnie testów pytań kontrolnych, przeprowadzenie testów POT, o ile było to tylko możliwe.

3. W 1959 roku profesor psychologii i psychiatrii na Uniwersytecie w Minnesocie, David Thorsten Lykken, opisał swoje eksperymenty¹⁵, rozpoczynając tym samym krytykę dotychczasowych technik badań poligraficznych. Trzeba przypomnieć, że Lykken nie zajmował się dotąd ani detekcją kłamstwa, ani psychologią sądową. Jego zainteresowania koncentrowały się na badaniach osobowości, w szczególności osobowości antysocjalnej, w których to dziedzinach był niekwestionowanym autorytetem. Do klasyki należy jego praca *A study of anxiety In the sociopathic personality*¹⁶, w której wskazał, że podstawowym objawem, a zarazem mechanizmem psychopatii (socioopatii), jest deficyt lęku. Badacz nie miał też żadnych doświadczeń śledczych. To ostatnie stało się zapewne jednym z ważniejszych powodów, dla których przyjął błędne dla swojej teorii założenia (por. niżej). Nie dysponował też wielokanałowym poligrafem. Z końcem lat 50. XX wieku standardowy poligraf był 3-kanałowy, zapisując przebieg czynności oddychania, zmiany w zakresie GSR oraz przebieg pracy układu krążenia. W swoim eksperymencie Lykken wykorzystał jedynie psychogalwanometr. Eksperyment ów, w którym uczestniczyli studenci, zgodnie z założeniami nie był eksperymentem mierzącym do wykrycia kłamstwa (lie) czy nieszczerości (deception), ale do wykrycia winy (detection of guilt). Później Lykken sprecyzował, że wykrywał też nie „winę” (guilt), ale „świadomość winnego” (guilty knowledge). Swoich uczestniczących w eksperymencie studentów podzielił na dwie grupy. Pierwszą stanowili „winni”, którzy uczestniczyli w wyreżyserowanym „przestępstwie”. Druga, „niewinni” – nieuczestniczący w zdarzeniu, ani nieznający jego szczegółów. Obydwie grupy badane były testami, które były modyfikacją testów POT. Badani nie mieli, jak to ma miejsce w badaniach poligraficznych, odpowiadać na pytania, a jedynie siedzieć nieruchomo i słuchać zadawanych im pytań. Aby ułatwić im koncentrację na pytaniach, mieli zasłonięte oczy, a na uszach słuchawki, przez które zadawane im były pytania. W drugiej wersji eksperymentu, zamiast milczeć, mieli powtarzać zadawane im pytania. Tak więc, w żadnej z wersji eksperymentu, nie odpowiadali oni na pytania, a zatem nie mieli okazji kłamać. Pytania zadawane były według schematu testu POT. Rozpoczęcie ich zadawania poprzedzone było wstępem eksperymentatora: „jeżeli ty jesteś «mordercą», będziesz wiedział co było niezwykłym przedmiotem, prezentowanym w pokoju, w którym dokonano «morderstwa»”¹⁷. Po tym wstępie następowało wyliczanie przedmiotów, z których jeden, sztaluga malarska, był wyeksponowany na miejscu zdarzenia. Eksperymentator pytał:

¹³ Por. J. Widacki, *Analiza przesłanek diagnozowania w badaniach poligraficznych*, Wydawnictwo UŚ, Katowice 1982.

¹⁴ Por. J. Reid, F. Inbau, *Truth and Deception. The Polygraph (Lie-detector) Technique*, William & Wilkins Comp, Baltimore 1977, s. 31.

¹⁵ D.T. Lykken, *The GSR In detection of guilt*, „Journal of Applied Psychology” 1959, 43, 6, s. 385–386.

¹⁶ „Journal of Abnormal and Social Psychology” 1957, 55, 1, s. 6–10.

¹⁷ D.T. Lykken, *The GSR In detection of guilt...*, op. cit.

- 1) Czy to był magnetofon?
- 2) Czy to była bombonierka?
- 3) Czy to były szachy?
- 4) Czy to była sztaluga?¹⁸

Dla wzmocnienia motywacji badanych, byli oni uprzedzani, że w przypadku wystąpienia reakcji na pytanie krytyczne, otrzymają niemiłe uderzenie prądem. W drugiej wersji, wyznaczono im nagrodę w wysokości 100 USD, co było kwotą dość znaczącą.

W pierwszym eksperymencie Lykken bezbłędnie wytypował 100% „niewinnych i aż 93,9% winnych. W drugim przypadku, w obydwu grupach, „winnych” i „niewinnych” trafnych wskazań miał 100%.

Były to wyniki lepsze niż te, które uzyskiwali eksperymetatorzy prowadzący eksperymentalne badania poligraficzne tradycyjnymi technikami, wyższe niż uzyskiwano w praktyce.

W oparciu o wyniki swych eksperymentów, Lykken dokonał surowej krytyki techniki klasycznej i techniki pytań kontrolnych (wówczas była to tylko technika Reida, w swej pierwotnej wersji). Uczony skrytykował same założenia techniki pytań kontrolnych. Uznał, że samo założenie porównywalności emocji wywołanej pytaniem krytycznym i kontrolnym jest, z teoretycznego punktu widzenia, nonsensem¹⁹. By miało sens – pytanie krytyczne i kontrolne musiałyby dotyczyć dwóch równie ciężkich przestępstw, co z oczywistych względów jest nie do zrealizowania.

Lykken krytykował też samą formę pytań zadawanych badanym w tradycyjnym badaniu poligraficznym, prowadzonym techniką klasyczną lub techniką pytań kontrolnych, takich jak „Czy to ty strzelałeś do Jonesa?” Jego zdaniem, reakcja osoby niewinnej (Ri) na tak zadane pytanie, wyglądałaby następująco:

$$Ri - LiFi \left(\frac{1 - Ci}{2} \right) + M1$$

Natomiast reakcja osoby winnej (Rg) wyglądałaby:

$$Rg = LgFg \left(\frac{1 - Cg}{2} \right) + M2$$

Gdzie:

M – jest średnią reakcją na pytania niezwiązane (obojętne);

L – to reaktywność autonomiczna (labilność) badanego;

F – to strach badanego przed konsekwencjami uznania za winnego;

C – przekonanie badanego o skuteczności testu, przy czym poziom tego zaufania uszeregowany jest od -1 (brak zaufania) do +1 (pełne zaufanie)²⁰.

Zgodnie z twierdzeniem Lykkena, jeśli osoba niewinnie podejrzana mówi prawdę i nie boi się, że będzie błędnie uznana za kłamcę, czyli faktycznie sprawcę przestępstwa, którego nie popełniła (wówczas $F=0$), a zatem:

$$Ri = 0 \left(\frac{1 - C}{2} \right) + M$$

¹⁸ *Ibidem*.

¹⁹ D.T. Lykken, *Psychology and lie-detection industry*, „American Psychologist” 1974, 29, 10, s. 732.

²⁰ *Ibidem*, s. 731.

Czyli $R_i = M$, a zatem reakcja na pytanie krytyczne będzie równa przeciętnej reakcji na pytanie obojętne. Jednak w sytuacji, gdy osoba niewinna (prawdomówna w badaniu) boi się badania, nie ma do niego pełnego zaufania, jej reakcja na pytanie krytyczne będzie silniejsza niż na pytanie obojętne. Wówczas wynik badania będzie taki, że zostanie ona uznana za osobę kłamiącą (nieszczerą)²¹. Co więcej Lykken podejrzewa, że niewinny i prawdomówny badany o dużej labilności emocjonalnej, może na pytania krytyczne zareagować silniej niż osoba winna i kłamiąca, ale o osobowości psychopatycznej²².

Wszystko to odnosiło się do techniki klasycznej. Równie surowo jednak Lykken ocenił technikę pytań kontrolnych. Jego zdaniem, pytania kontrolne w takiej formie, jak to jest przyjęte w tej technice, nie mogą być standardem, bowiem ich waga emocjonalna nie jest adekwatna do wagi pytań krytycznych. Wedle założeń techniki pytań kontrolnych, spodziewany „rachunek kłamstwa” (lie-score) wygląda następująco:

$$R_x - R_y = LF \left(\frac{1 - C}{2} \right) + M + LF \left(\frac{1 - C}{2} \right) + M = 0$$

Gdzie:

R_x – jest reakcją na pytanie krytyczne;

R_y – jest reakcją na pytanie kontrolne.

Natomiast u osoby niewinnej i prawdomównej wyglądałoby to następująco:

$$R_x - R_y = LF \left(\frac{1 + C}{2} \right) + M - LF \left(\frac{1 - C}{2} \right) + M = CLF$$

Tak by było, twierdzi Lykken, gdyby waga pytań kontrolnych odpowiadała wadze pytań krytycznych. Jednak z samej istoty rzeczy tak nie jest, zatem założenia teoretyczne techniki pytań kontrolnych są błędne²³.

Zdaniem Lykkena, jedyną poprawną metodologicznie i zgodną z wiedzą psychologiczną techniką badania poligraficznego jest technika zaproponowana przez niego. Owa technika rezygnuje z używania pytań, na które badany ma odpowiadać „tak” lub „nie”, przy założeniu, że kłamca/sprawca, odpowiadając na pytanie krytyczne słowem „nie”, będzie kłamał, a jego kłamstwo wykryć ma poligrafer. W technice Lykkena badany w ogóle nie odpowiada „tak” ani „nie”, nie ma więc okazji kłamać. To, co wykrywa technika Lykkena, to nie kłamstwo, ale „świadomość winnego” (guilty knowledge). Ta „świadomość winnego”, czyli znajomość szczegółów zdarzenia, które jest przedmiotem badania, różnicuje sprawców i nie-sprawców.

Posługując się językiem kryminalistyki można by powiedzieć, że zaproponowana przez Lykkena technika, jest techniką polegającą na obserwacji zmian emocjonalnych, które pozwalają wnioskować o istnieniu, bądź nieistnieniu u badanego pamięciowych śladów zdarzenia.

Konstrukcja testów w tej technice jest zasadniczo identyczna, jak znanych i stosowanych pomocniczo w ramach techniki klasycznej i techniki pytań kontrolnych testów szczytowego napięcia (Peak of Tension, POT).

Lykken propozycję swej techniki przedstawia na fikcyjnym przykładzie. Załóżmy, pisze Lykken, że dokonano napadu na jedną z kas pożyczkowych na Manhattanie, nazywającą się „Friendly Loan Company”. Sprawca podszedł do okienka, powiedział urzędnicze, że potrzebuje pożyczki, bowiem

²¹ *Ibidem*.

²² *Ibidem*.

²³ *Ibidem*, s. 732.

są mu potrzebne pieniądze na wysokie honorarium dla lekarza. W czasie rozmowy wyjął pistolet, zrabował pieniądze i zbiegł.

W takiej sprawie, zdaniem Lykkena, badanie podejrzanego powinno wyglądać tak, że przygotowano by kilka zestawów pytań. Przed ich zadaniem, poligrafer wygłaszałby wstęp: „Dokonano napadu na kasę pożyczkową na Manhattanie. Jeżeli to ty dokonałeś napadu, to znasz nazwę tej kasy. Teraz siedź spokojnie i powtarzaj za mną

- a) Czy to było Ideal Loan Company?
- b) Czy to było Continental Loan Company?
- c) Czy to było Guarante Loan Company?
- d) Czy to było Friendly Loan Company?
- e) Czy to było Fidelity Loan Company?”

Sprawca napadu ma świadomość, że napadu dokonano na Friendly Loan Company. To jest właśnie ta „świadomość winnego” (guilty knowledge), której siłą rzeczy nie ma osoba niewinna. Tę świadomość można wykryć badaniem poligraficznym, a badany wcale nie musi przy tym werbalnie kłamać. Prawdopodobieństwo, że badany przypadkowo zareaguje na właściwe pytanie, przy zadanych pięciu pytaniach, wynosi jak 1:5, czyli 20%, dlatego należy przeprowadzać następne testy. W następnym teście można na przykład więc pytać, co napastnik powiedział urzędniczce przed napadem, na co potrzebne mu są pieniądze.

Przypadkowa reakcja na właściwe pytanie w dwóch testach, z których każdy zawiera 5 pytań, wynosi $0,2 \times 0,2 = 0,04$. Czyli teoretycznie, osoba niewinna ma szanse przypadkowo zareagować na właściwe pytanie w dwóch kolejnych testach z prawdopodobieństwem 0,04, czyli w 4 przypadkach na 100.

Gdyby przeprowadzić trzy testy, prawdopodobieństwo przypadkowej reakcji we wszystkich trzech testach wynosiłoby $0,2 \times 0,2 \times 0,2 = 0,008$. Czyli w 8 przypadkach na 1000. W przypadku przeprowadzenia 10 testów (każdy po 5 pytań), jak łatwo policzyć przypadkowa reakcja na wszystkie pytania krytyczne wynosiłaby 1:1 000 000.

Zatem prawdopodobieństwo przypadkowego zareagowania przez osobę niewinną, niemającą „świadomości winnego”, można obliczyć wedle poniższego wzoru:

$$\left(\frac{1}{K}\right)^N$$

Gdzie:

K – to liczba alternatywnych możliwości (liczba pytań w teście)

N – to liczba przeprowadzonych testów.

Swoje przemyślenia związane z własną, nową techniką badań poligraficznych, opierającą się na testach Guilty Knowledge Test (GKT; której nazwę w Polsce tłumaczono jako „technika wiedzy o czynie”), krytykę dotychczas stosowanych technik badań (klasyczną i pytań kontrolnych), a także krytykę amerykańskiego „przemysłu lie-detection”, Lykken zebrał w swojej książce *A tremor In the blood*, wydanej w roku 1981²⁴.

4. Proponowana przez Lykkena technika badań poligraficznych była nowością, jakkolwiek jednak jej podstawą były, nieco zmodyfikowane, testy szczytowego napięcia (Peak of Tension, POT). Testy tego typu znane były i stosowane w praktyce od lat 30. XX wieku, jako uzupełnienie technik klasycznej i pytań kontrolnych. Jak już wspomniano, oparcie techniki badań wyłącznie na testach POT postulo-

²⁴ D.T. Lykken, *The tremor In the blood*, McGraw-Hill Book Comp., New York, St. Louis, San Francisco (et All.) 1981.

wał jeszcze w okresie międzywojennym (Lee w liście do szefa FBI Edgara Hoovera)²⁵. W latach 50. XX wieku, podobny postulat zgłaszał Burack²⁶. Lykken nie był więc pierwszym, który sugerował oparcie badania na testach szczytowego napięcia (POT). Co więcej, jeszcze przed Lykkenem stosowano także w technikach pytań kontrolnych testy, na które badany albo w ogóle nic nie odpowiadał (tzw. Silent Answer Test, „test milczących odpowiedzi), albo bez względu na to, co było treścią pytań, miał odpowiadać „tak” („Yes-test”).

Krytycy Lykkena podnosili także, że bazę empiryczną nowej techniki stanowią zaledwie dwa eksperymenty ze studentami, których reakcje obserwowano jedynie przy pomocy psychogalwanometru²⁷. Powtórzenie eksperymentu Lykkena²⁸ nie potwierdziło tak wysokiej wartości diagnostycznej proponowanej techniki. Uzyskano jedynie 77,7% trafnych wskazań, a więc porównywalnie, a nawet mniej niż uzyskiwali eksperymetatorzy posługujący się techniką pytań kontrolnych.

Podnoszono także, że technika ta jest wysoce niepraktyczna, bowiem w realnej sprawie karnej trudno znaleźć kilka szczegółów, których znajomość różnicuje sprawców i nie-sprawców, a już na pewno ich znajomość nie różnicuje sprawców i świadków²⁹.

W praktyce polskiej lat 70. i 80. XX wieku, kiedy powszechnie stosowano technikę Reida (technikę pytań kontrolnych Reida) i zaczęto stosować technikę Backstera, zawsze kiedy tylko było to możliwe, obok testów pytań kontrolnych, wykonywano testy szczytowego napięcia (POT). Tak było w blisko 80% przeprowadzonych badań. Jednak nigdy nie udawało się skonstruować więcej niż dwa testy POT. Oznacza to, że nigdy w praktyce nie udało się odnaleźć więcej niż dwóch szczegółów zdarzenia, które w przeciwieństwie do osoby niewinnej, sprawca powinien był spostrzec i zapamiętać³⁰.

Tę niepraktyczność techniki Lykkena, wykazał później Podlesny, który po przeanalizowaniu 758 badań poligraficznych wykonanych przez FBI stwierdził, że w zaledwie 51 przypadkach (6,7%) teoretycznie badanie tą techniką było możliwe, to znaczy, że jedynie w tak niewielu sprawach spełnione były warunki do przeprowadzenia badań techniką GKT³¹. W sytuacji, w której założono by, zgodnie z sugestią Lykkena, że trzeba przeprowadzić aż 6 testów (inaczej mówiąc: gdyby chcieć znaleźć 6 szczegółów, których znajomość różnicuje sprawców od nie-sprawców), teoretycznie byłoby to możliwe tylko w 16 przypadkach (2,1%). Co więcej okazało się, że badania tą techniką dają średnio 5% wskazań pozytywnie fałszywych (uznanie 5% osób niewinnych za winnych) i aż 20% wskazań fałszywie negatywnych, czyli uznanie aż 20% winnych, mających „świadomość winnego” za niewinnych. Ten relatywnie wysoki, bo 20% wynik wskazań fałszywie negatywnych, świadczy o tym, że każdorazowo na 100 winnych, 20 osób winnych uznanych zostanie za niewinne.

Krótko mówiąc, technika Lykkena (Guilty Knowledge Technique) okazała się techniką nie dość, że bardzo niepraktyczną, to jeszcze wykazującą wartość diagnostyczną nie lepszą od techniki pytań kontrolnych. Rodzić się może pytanie, dlaczego mimo takiej poprawności teoretycznej i metodologicznej, technika „Guilty Knowledge Technique” nie okazała się lepsza od technik pytań kontrolnych?

Wydaje się, że Lykken przyjął błędne założenie, że sprawca przestępstwa, zwłaszcza zabójstwa (w tych sprawach najczęściej wykorzystuje się badania poligraficzne), spostrzega i zapamiętuje

²⁵ Por. D. Krapohl, P. Shaw, *Fundamentals of polygraph practice*, Elsevier, Academic Press, 2015, s. 27.

²⁶ B. Burack, *A critical analysis of the theory and limitation of the lie-detector*, „Journal of Criminal Law, Criminology and Police Science” 1955, 46, 3, s. 414–426.

²⁷ G. Ben Shakhar, I. Lieblich, S. Kugelmass, *Guilty Knowledge Technique: application of Signac detection measures*, „Journal of Applied Psychology” 1970, 54, 5, s. 409–413.

²⁸ *Ibidem*, s. 410.

²⁹ Por. J. Konieczny, M. Frąś, J. Widacki, *Pochodzenie ukrytej informacji a niektóre cechy osobowości w badaniu poligraficznym*, *Archiwum Medycyny Sądowej i Kryminologii* 1984, 34, 1; por. także: M. Tarabula, M. Widacki, *The Mount of information remembered by the perpetrator in the context of application of the GKT in criminal investigation*, „European Polygraph” 2016, 10, 2 (36), s. 63–76.

³⁰ Por. A. Krzyścin, *Badania poligraficzne wykonane techniką Reida- analiza doświadczeń polskich*, Katowice 1980 (niepublikowana praca doktorska); por. także: J. Widacki, *Analiza przesłanek diagnozowania w badaniach poligraficznych*, Wydawnictwo Uniwersytetu Śląskiego, Katowice 1982.

³¹ J.A. Podlesny, *A paucity of operable case facto restricts applicability of the Guilty Knowledge Technique In FBI criminal polygraph examinations*, *Forensic Sciences Communications*, 2003, 5, 3.

wszystkie szczegóły zdarzenia. Otóż, jak pokazuje praktyka, a potwierdzają badania naukowe³², sytuacja emocjonalna sprawcy w chwili popełniania przestępstwa, szczególnie zabójstwa, jest tak szczególna, że dochodzi do zwężenia świadomości (oczywiście nie ma to nic wspólnego z poczytalnością!), że nie spostrzega lub nie zapamiętuje wielu szczegółów, co do których jesteśmy przekonani, że sprawca powinien zapamiętać. Pytany o te niezapamiętane szczegóły, sprawca nie reaguje, podobnie jak nie reaguje osoba niewinna.

Dla przykładu można dodać, że w kilku ostatnich sprawach o zabójstwo, z praktyki jednej tylko kancelarii adwokackiej – zabójcy, którzy przyznali się do przestępstw, nie pamiętali jak ubrana była ofiara, którą zaatakowali z bliska, nie pamiętali nawet w przybliżeniu liczby ciosów, jakie ofierze zadali³³. W pierwszym przypadku sprawca, który faktycznie zadał ponad 40 ciosów nożem, był przekonany, że zadał ich nie więcej niż 3–4. W drugim przypadku sprawca zadał ofierze ponad 20 ciosów sądząc, że zadał ich nie więcej niż 4. Gdyby ci zabójcy nie przyznali się do zabójstw, a byli badani poligraficznie techniką GKT i byli pytani o ubranie ofiary lub liczbę zadanych ciosów, zareagowałiby tak, jak osoby niewinne.

Dlatego też do równania Lykken'a:

$$\left(\frac{1}{K}\right)^N$$

należałoby dopisać warunek: „jeżeli sprawca pamięta szczegóły, o które jest pytany w testach GKT, to...”, czyli równanie to powinno wyglądać tak:

$$\rightarrow \left(\frac{1}{K}\right)^N 1 \wedge p \rightarrow \left(\frac{1}{K}\right)^N$$

(p – sprawca pamięta te szczegóły, o które jest pytany)

Przy okazji rodzi się pytanie, którego Lykken w swych rozważaniach nie przewidział. Co wtedy, gdy w 6. przeprowadzonych testach GKT badany zareaguje na pytanie krytyczne (pytanie „klucz”) na przykład tylko w 3 testach, a w trzech dalszych nie zareaguje? Jeśli idzie o krytyczne uwagi Lykken'a, zgłoszone do techniki pytań kontrolnych, to twierdzenie, że z założenia pytanie kontrolne nie może spełnić przypisanej mu roli, bowiem jego waga emocjonalna zawsze będzie mniejsza od wagi pytania kontrolnego, nie potwierdza się w tysiącach badań wykonywanych w praktyce na całym świecie, ani też w licznych badaniach eksperymentalnych. Jedne i drugie potwierdzają relatywnie wysoką wartość diagnostyczną badań poligraficznych wykonywanych technikami pytań kontrolnych.

Natomiast uwagi Lykken'a, dotyczące warunków poprawności testu pytań kontrolnych, są generalnie słuszne. Rzecz w tym, że Lykken uważa, iż warunki te są nie do spełnienia, a praktyka na całym świecie pokazuje, że można je spełnić. Nie zmienia to faktu, że Lykken świadomie czy nie, zwrócił uwagę na warunki poprawności badania poligraficznego techniką pytań kontrolnych. Aby to badanie miało sens, spełnione jednak muszą być pewne warunki. W szczególności osoba badana musi mieć zaufanie do skuteczności badania poligraficznego; musi mieć przekonanie, że badanie to rzeczywiście jest w stanie wykryć kłamstwo lub potwierdzić prawdomówność. Spełnienie tego warunku w dużej mierze zależy od tego, czy prawidłowo został przeprowadzony wywiad przedtestowy, czy właściwie dobrano pytanie kontrolne i czy należycie zostało ono w wywiadzie omówione. Kolejnym warunkiem jest prawidłowo przeprowadzony test stymulacyjny (np. test z cyfrą).

³² Por. S.A. Christansen, I. Frej, E. v. Kremnitzer, *Searching for offender's memories of violent crime*, [w:] *Offender's memories of violent crime*, ed. S.A. Christianson, John Wiley and Sons Ltd., 2007.

³³ Sprawy: sygn. PO.I Ds.43.2016, Prokuratura Okręgowa w Krakowie; sygn. PO.I Ds., sprawa sygn. III K 337/02, Sąd Okręgowy w Krakowie; sprawa sygn. II I K 61/14, Sąd Okręgowy w Nowym Sączu.

Jednakże uwagi Lykkena o wpływie osobowości badanego na sposób reagowania, w szczególności twierdzenie, że na pytania krytyczne w teście pytań kontrolnych labilny emocjonalnie niewinny – zareaguje jak winny, psychopatyczny winny – nie zareaguje, ergo – zareaguje jak osoba niewinna, nie doczekały się potwierdzenia ani w praktyce, ani w badaniach eksperymentalnych. Badania Barlanda i Raskina³⁴ nie wykazały różnic w reakcjach poligraficznych osób o najwyższym i najniższym wskaźniku psychopatii, mierzonym testem MMPI. Również inni autorzy, w tym Larson³⁵, Reid i Inbau³⁶ uważali, że psychopaci nie są gorszymi podmiotami badań poligraficznych niż nie-psychopaci.

Jeśli idzie o osoby neurotyczne (labilne emocjonalnie), to zdaniem Dufka³⁷, są najlepszymi obiektami badań poligraficznych. Spotkać też można pogląd, że poza nerwicą natręctw, żadna inna nie stanowi przeszkody w skutecznym badaniu poligraficznym³⁸.

Abstract

Polygraph testing method controversy

Beginning with the 1920s polygraph examinations were conducted in the so-called „classical technique”, making use of two types of questions: critical – connected to the subject of the investigation, and background – unconnected to the issue at question and therefore indifferent. John Reid’s Control Questions Technique (CQT) was a new technique of examination, introduced in 1947. Reid added a third type of questions, namely control questions, to polygraph testing. They were not connected to the investigated issue, yet they referred to matters that were similar in type to the ones addressed by critical questions. Control questions were formulated so that the examinee answered them „no”, and saying that either purposefully lied or was not sure whether he lied or not. The technique developed into a large family of control question techniques (Backster’s technique and others based on it). In the 1960s, David Lykken questioned the methodological correctness of both classical and control questions techniques. In their place, he proposed his technique of examination, which he later called the Guilty Knowledge Technique (GKT). The charges put forward by Lykken were challenged, and his technique had two fundamental charges levelled against: first, that it follows an erroneous premise that the perpetrator remembers a great deal of detail from the moment of crime and that they can be elicited during the examination, while the other concerned its high impracticality. Nonetheless, the dispute with Lykken, whose charges were mostly rejected, made an impact as it improved the methodology of the control question techniques.

Key words: polygraph testing techniques, control questions techniques, Lykken technique

³⁴ G. Barland, D. Raskin, *Psychopathy and detection of Deception In criminal suspects*, Utah Polygraph Association Newsletter 1974, 1, 11.

³⁵ J. Larson, *The cardio-pneumo psychogram and its use In the emotions with practical applications*, „Journal of Experimental Psychology” 1922, 5, s. 323–328.

³⁶ J. Reid, F. Inbau, *Truth and Deception. The polygraph (lie-detector) technique*, William & Wilkins, Baltimore 1977, s. 201.

³⁷ M. Dufek, *K problematycie polygrafického vysetrovani v kriminalistice*, [w:] *Doplňkové studijní materiály pro kriminalistický smer prvnického studia*, Praha 1970, s. 70–71.

³⁸ Por. L. Wdowiak, R. Rutkowski, *Wpływ stanów psychopatologicznych na reakcje emocjonalno-wegetatywne człowieka w aspekcie badań wariograficznych*, Materiały Sympozjum, Toruń 1976, s. 14.

Część III

Subiektywizm w badaniach poligraficznych

Subiektywizm, a ściślej margines subiektywizmu w badaniach poligraficznych jest istotnym problemem zarówno naukowym, czysto poznawczym, jak i praktycznym. Wielkość tego marginesu jest jednym z podstawowych kryteriów pozwalających na uznanie jakiejś metody za dostarczającą dowodu naukowego (*scientific evidence*) w procesie. To właśnie szeroki margines subiektywizmu nie pozwala na zaliczenie ekspertyzy pismoznawczej (identyfikacji pisma ręcznego) do metod dostarczających dowodu naukowego, takiego jakim jest na przykład dowód z badań DNA.

W badaniu poligraficznym badający ekspert ocenia reakcje badanego zapisane przez poligraf. Każda ocena, w przeciwieństwie do pomiaru, niesie ze sobą ryzyko subiektywizmu.

Źródeł subiektywizmu interpretacji zapisów poligraficznych jest wiele. Składać się na niego może motywacja badającego, na którą z kolei mogą mieć wpływ rozmaite czynniki (w tym przekonanie o ewentualnej winie badanego wynikające z innych, zebranych już dowodów czy na przykład sugestii płynących od zlecającego badanie organu procesowego), właściwości osobnicze badającego (w tym jego podatność na sugestię), jego aktualna predyspozycja psychofizyczna, stopień wykształcenia, doświadczenie zawodowe itp.

Jednym ze sposobów ograniczenia marginesu subiektywizmu w badaniach poligraficznych jest wprowadzenie numerycznych metod interpretacji zapisów poligraficznych. Metoda numeryczna narzuca pewne ramy ocen, tym samym ogranicza ich arbitralność.

Przyjmując to założenie jako ogólną hipotezę, postanowiliśmy w oparciu o materiał uzyskany z naszych badań eksperymentalnych zweryfikować trzy hipotezy szczegółowe:

- 1) Metody numeryczne pozostawiają mniejszy margines subiektywizmu niż metody jakościowe (u nas eufemistycznie zwane „globalnymi” lub „holistycznymi”);
- 2) Metody o mniejszej skali (np. 2- lub 3-stopniowe) są bardziej obiektywne niż 7-stopniowe;
- 3) Oceniający zapisy metodą „ślepej interpretacji” (poprzez wyeliminowanie innych niż czysty zapis czynników, na których opiera diagnozę) uzyskuje większy odsetek trafnych diagnoz.

Sprawdzanie tych hipotez będzie przedmiotem rozprawy doktorskiej jednego z uczestników projektu (Marcina Gołaszewskiego). Ogólna problematyka subiektywizmu w badaniach poligraficznych była też przedmiotem studium opublikowanego w „Studiach Prawniczych” (2017, nr 2) – por. Aneks C.

ANEKS C

Jan Widacki, Marcin Gołaszewski

Subiektywizm w badaniach poligraficznych*

Wprowadzenie

Problem subiektywizmu w badaniach naukowych jest interesujący dla metodologii ogólnej nauk, a także dla poszczególnych dyscyplin naukowych. W metodologii badań, subiektywizm oznacza margines swobody interpretacji badacza, niepodlegający obiektywnym kryteriom, a raczej osobistym przekonaniom badacza. Innymi słowy, subiektywizm to pierwiastek mający wpływ na wynik badania, zależący nie od przedmiotu badanego, obiektywnie istniejącego i obiektywnie posiadającego jakieś cechy, także nie od przyrządu pomiarowego, ale od badacza, czyli podmiotu poznającego, jego właściwości i cech, a czasem także przekonań i przyjętych *a priori*, bez koniecznego uzasadnienia założeń. Problem subiektywizmu dotyczy wszystkich nauk, od historycznych poczynając¹, poprzez nauki społeczne, aż po nauki przyrodnicze.

Wydaje się, że im dyscyplina bardziej ścisła, im bardziej zaawansowana metodologicznie, tym mniejszy w niej margines subiektywizmu. Zgodnie z tym założeniem, margines subiektywizmu z zasady będzie większy w socjologii czy psychologii niż w fizyce czy chemii. W naukach społecznych, jak się zdaje, akceptowany przez badacza pogląd – a tym bardziej przyjęta przez niego teoria – mają wpływ zarówno na stawianie problemów, budowanie hipotez, jak i później na interpretację wyników badań empirycznych, które te hipotezy sprawdzają². W naukach przyrodniczych, takich jak fizyka czy chemia, jak wspomniano, margines subiektywizmu w badaniach wydaje się z natury rzeczy mniejszy – w zasadzie daje się sprowadzić do błędu pomiaru lub jego błędnej interpretacji. W metodologii fizyki rozróżnia się „błędy pomiarowe” (błąd przybliżenia, błąd przeoczenia, pomyłki) oraz „niepewności pomiarowe” (niepewność wzorcowania, niepewność eksperymentatora, niepewność przypadkowa)³. Szczególnie w „niepewności eksperymentatora” kryć się może pierwiastek subiektywizmu. Problem subiektywizmu w filozofii nauki i metodologii jest też rozważany przy okazji porównywania metod ilościowych i jakościowych⁴.

* J. Widacki, M. Gołaszewski, *Subiektywizm w badaniach poligraficznych*, „Studia Prawnicze. Rozprawy i Materiały” 2017, nr 2 (21), s. 187–214.

¹ Artykuł powstał w ramach projektu DEC-2013/11/B/HS5/03856 finansowanego ze środków Narodowego Centrum Nauki. Por. B. Miśkiewicz, *Wprowadzenie do badań historycznych*, Poznań 1993, s. 165; J. Topolski, *Jak się pisze i rozumie historię. Tajemnice narracji historycznej*, Warszawa 1996, s. 339.

² Por. J. Czapiński, *Wartościowanie – zjawisko inklinacji pozytywnej. O naturze optymizmu*, Wrocław 1985; M. Materska, *Psychologiczna i formalna analiza sądów oceniających*, [w:] *Psychologia a poznanie*, red. M. Materska, T. Tyszka, Warszawa 1997; T. Tyszka, *Psychologiczne pułapki podejmowania decyzji*, Gdańsk 2000; M. Weber, *Obiektywność poznania w naukach społecznych*, [w:] *Problemy socjologii wiedzy*, Warszawa 1985.

³ Por. H. Szydłowski, *Pracownia fizyczna wspomagana komputerowo*, Warszawa 2012.

⁴ Por. np. R.G. Long, M.C. White, W.H. Friedman, D.V. Brazeal, *The „Qualitative” versus „Quantitative” research Debate: A Question of Metaphorical Assumptions?*, „International Journal of Value-Based Management” 2000, t. 13, nr 2, s. 189–197.

Subiektywizm w naukach sądowych

O ile problem subiektywizmu badacza w naukach podstawowych jest interesujący z metodologicznego, czy nawet filozoficznego punktu widzenia, a w badaniach w naukach przyrodniczych zbyt szeroki margines subiektywizmu prowadzić może do co najwyżej błędnych teorii, to w naukach sądowych (ang. *forensic sciences*) ma on sens jak najbardziej praktyczny – co więcej często istotny dla losów konkretnego człowieka, przeciwko któremu toczy się postępowanie sądowe. W ekspertyzie mamy do czynienia z badaniem (wyjaśnianiem) faktu jednostkowego, w oparciu o ustalenia jakiejś dyscypliny naukowej. Jak zauważył kiedyś Józef Życiński⁵:

W procedurach badawczych występujących przy klasyfikacji i wyjaśnianiu faktów jednostkowych nie ma role odgrywają elementy pozadyskursywne uzależnione od indywidualnych predyspozycji poszczególnych badaczy, ich intuicji, zdolności kojarzenia faktów, wcześniejszych doświadczeń itp. [...] Przecenianie roli czynników pozadyskursywnych w poznaniu może jednak prowadzić łatwo do subiektywizmu, w którym przez odwołanie do osobistych długoletnich doświadczeń usiłowano by wprowadzać interpretacje niemożliwe do racjonalnego uzasadnienia⁶.

Ekspertyzy w naukach sądowych opierają się na opisach, pomiarach i ocenach. Nawet pomiar będący istotą wielu ekspertyz w naukach sądowych obciążony jest, jak każdy pomiar, pewną niedokładnością czy niepewnością (ang. *uncertainty*), która może rzutować na wartość wydanej w oparciu o ten pomiar opinii. Może być przyczyną błędu lub pomyłki, może też rodzić niepewność⁷ co do wiarygodności opinii.

Niedokładność pomiaru może być wynikiem niedoskonałości narzędzia pomiarowego lub sposobu jego użycia. Stąd też konstruuje się coraz dokładniejsze narzędzia pomiarowe. Niewłaściwy sposób użycia narzędzi pomiarowych może wynikać z niefachowości osoby dokonującej pomiaru, jej niezręczności czy niedbalstwa. Czasem z innych jeszcze powodów, jak na przykład wymuszony pośpiech czy szczególne, niesprzyjające okoliczności dokonywania pomiaru. Dotyczy to mierzenia, ważenia lub porównywania uzyskanych wyników ilościowych. Ma to znaczenie dla wszystkich ekspertyz opartych na metodach ilościowych, a więc przede wszystkim w chemii sądowej, toksykologii sądowej, biologii sądowej i podobnych dyscyplinach. W większości ekspertyz wielkość błędu pomiaru w pewnych określonych granicach (por. niżej) jest dopuszczalna i nie ma wpływu na ostateczny wynik i oparte na nim wnioski.

Nauka posługuje się między innymi pojęciem błędu standardowego (lub błędu średniego), przez który rozumie się odchylenie średnie wyników pomiarów tej samej wielkości, które otrzymano przy użyciu tego samego narzędzia pomiarowego⁸.

Dokonanie pomiaru (lub opisu) przy ekspertyzie jest zwykle dopiero punktem wyjścia do dalszych działań, których finałem jest wydanie opinii. Opinia bowiem to – jak słusznie zauważył kiedyś Kazimierz Jaegermann – na pewno coś więcej niż sam wynik badania (opis czy pomiar)⁹. Wedle tego autora, na opinię składają się działania pomiarowo-opisowe, działania interpretacyjne i wnioskowanie¹⁰. Opinia końcowa (wnioski z opinii) jest zaś tak naprawdę decyzją biegłego¹¹.

⁵ J. Życiński, *Problem wyjaśniania w naukach nomotetycznych a poznanie faktu jednostkowego*, [w:] *Z zagadnień teorii opinii biegłego*, red. J. Widacki, Materiały IV Sympozjum Metodologii Kryminalistyki i Nauk Pokrewnych, Chęciny 24–25 VI 1983, Katowice 1983, s. 60–61.

⁶ *Ibidem*, s. 61.

⁷ S. Bell, *Measurement uncertainty in forensic sciences. A practical guide*, London–New York 2017, s. XIX.

⁸ Ten błąd standardowy (S) wyraża się wzorem:

$$\bar{s}_x = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n(n-1)}}$$

X_i – wartość uzyskana w pomiarze numer i , \bar{x} – wartość średnia wyników pomiarów, n – liczba pomiarów.

⁹ K. Jaegermann, *Opiniowanie sądowo-lekarskie. Eseje o teorii*, Warszawa 1991, s. 28.

¹⁰ *Ibidem*, s. 29.

¹¹ *Ibidem*, s. 33.

Błąd pomiaru nabiera szczególnego znaczenia, gdy przepis prawa lub orzecznictwo sądowe wprowadza progi ilościowe istotne dla kwalifikacji prawnej czynu. Tak na przykład zgodnie z art. 115 § 16 kodeksu karnego¹² stan nietrzeźwości zachodzi, gdy zawartość alkoholu we krwi przekracza 0,5 promila albo prowadzi do stężenia przekraczającego tę zawartość, lub gdy zawartość alkoholu w 1 dm³ wydychanego powietrza przekracza 0,25 mg albo prowadzi do stężenia przekraczającego tę wartość. Natomiast zawartość alkoholu w wydychanym powietrzu między 0,1 a 0,25 mg na 1 dm³ lub od 0,2 do 0,5 promila we krwi oznacza „stan po spożyciu alkoholu” (por. art. 46 ust. 2 ustawy o wychowaniu w trzeźwości¹³ oraz art. 87 § 1 kodeksu wykroczeń¹⁴). Podobnie jest z rozdziałem na „znaczną” bądź „nieznaczną” ilość środka psychotropowego (por. art. 62a ustawy o przeciwdziałaniu narkomanii¹⁵).

Jedną z pierwszych polskich prób sprawdzenia poprawności pomiarów zawartości alkoholu w krwi dokonywanych przez profesjonalne laboratoria, należące do instytucji państwowych (w tym czasie nie było jeszcze prywatnych laboratoriów), a więc laboratoriów akademickich, laboratoriów służby zdrowia i laboratoriów milicyjnych (które rutynowo wykonywały wówczas takie badania dla potrzeb organów ścigania i wymiaru sprawiedliwości, a te przyjmowały ich wyniki za podstawę swych orzeczeń), były badania przeprowadzone przez Instytut Ekspertyz Sądowych w 1966 r.¹⁶. Trzydzieści takich laboratoriów (pracowni) otrzymało pięć wzorcowych roztworów alkoholowych sporządzonych sztucznie, w warunkach laboratoryjnych w surowicy krwi ludzkiej. Jedna z próbek zawierała 0,17 promila alkoholu (tj. poniżej górnej granicy przyjętej dla alkoholu fizjologicznego), inna zawierała 0,23 promila alkoholu, czyli już powyżej granicy „stanu wskazującego na użycie alkoholu”, a inna – 0,48 promila alkoholu, czyli poniżej „progu nietrzeźwości” (który wynosił wtedy 0,50 promila alkoholu). Okazało się, że próbkę zawierającą 0,17 promila jedno z laboratoriów oceniło na 0,28 promila (czyli powyżej granicy „stanu wskazującego na użycie alkoholu”), próbki zawierające 0,23 promila alkoholu oceniano od 0,0 promila aż do 0,36 promila. Probki zawierające 0,48 promila oceniano od 0,22 promila (czyli poniżej granicy „stanu wskazującego na użycie alkoholu”), aż po 0,57 promila (powyżej granicy „stanu nietrzeźwości”). Trzeba pamiętać, że z każdym z tych bezkrytycznie przyjętych wyników wiążą się określone konsekwencje prawne. Niektóre osoby trzeźwe byłyby uznane za nietrzeźwe, niektóre nietrzeźwe – za trzeźwe. Pokazuje to przy okazji, jak ważna jest certyfikacja i atestacja laboratoriów, którego to problemu sądy i prokuratury zupełnie nie doceniają. W czasie gdy przeprowadzano opisywane badania certyfikacji i atestacji laboratoriów jeszcze nie było. Dziś realizowanej w podobny sposób procedurze poddają się dobrowolnie niektóre placówki badawcze z zakresu nauk sądowych. Jednak brak aktualnego atestu, a nawet niepowodzenie przy próbie jego uzyskania niestety w oczach organów ścigania i sądów nie dyskwalifikują takiego laboratorium i wyników jego badań, które są uznawane za dowody.

We wszystkich metodach ilościowych w naukach sądowych każda interpretacja pomiaru bezwzględnie musi brać pod uwagę fakt znanego i określonego błędu przy tego rodzaju pomiarach. Problem subiektywizmu w naukach sądowych zaczyna się wszędzie tam, gdzie wyniki pomiaru (bardziej lub mniej dokładnego) wymagają jeszcze oceny i interpretacji¹⁷.

Jest jednak szereg nauk sądowych, w których w ramach ekspertyzy nie dokonuje się w ogóle pomiarów fizycznych, a poprzestaje na opisie pewnych cech. Są też takie, w których pomiar ma jedynie

¹² Ustawa z dnia 6 czerwca 1997 r. – Kodeks karny, Dz.U. z 1997 r., nr 88, poz. 553.

¹³ Ustawa z dnia 26 października 1982 r. o wychowaniu w trzeźwości i przeciwdziałaniu alkoholizmowi, Dz.U. z 1982 r., nr 35, poz. 230.

¹⁴ Ustawa z dnia 20 maja 1971 r. – Kodeks wykroczeń, Dz.U. z 1971 r., nr 12, poz. 114.

¹⁵ Ustawa z dnia 29 lipca 2005 r. o przeciwdziałaniu narkomanii, Dz.U. z 2005 r., nr 179, poz. 1485.

¹⁶ J. Markiewicz, J. Nedoma, H. Serda, *Z badań nad dokładnością oznaczeń alkoholu we krwi*, „Problemy Kryminalistyki” 1966, nr 61–62, 1966, s. 472–478.

¹⁷ Por. Ch. O'Hara, J.W. Osterburg, *An introduction to criminalistics. The application of the physical sciences to the detection of crime*, Bloomington–London, 1972, s. 680 i n.; T.F. Kiely, *Forensic evidence: science and the criminal law*, Boca Raton–London–New York–Washington DC 2001, s. 2 i n.; T. Vosk, *Measurement uncertainty*, [w:] *Encyclopedia of Forensic Sciences*, red. J.A. Siegel, P.J. Saukko, Oxford 2013, s. 322–331; C.G.G. Aitken, F. Taroni, A. Biedermann, *Statistical interpretation of evidence: Bayesian analysis*, [w:] *Encyclopedia...*, s. 292–297; B. Robertson, G.A. Vignaux, Ch. Berger, *Interpreting evidence. Evaluating forensic sciences in the courtroom*, wyd. 2, Oxford 2016, s. 62–63, 102–104.

znaczenie drugorzędne, pomocnicze. Podstawą ustaleń poprzedzających wydanie opinii w takich dyscyplinach jest z zasady ocena dokonywana przez eksperta. Tak jest na przykład w psychologii sądowej, psychiatrii sądowej, klasycznej medycynie sądowej czy grafologii sądowej – i oczywiście w detekcji kłamstwa, zarówno instrumentalnej, jak i nieinstrumentalnej.

Ta ocena eksperta dokonywana jest w oparciu o pewne przyjęte w tej dyscyplinie kryteria. Niewątpliwie podstawowym jest określenie „typowości przypadku”. Biegły ocenia badany przez siebie przypadek, porównując go z przypadkiem typowym. Ten ostatni z samej istoty ma charakter statystyczny. „Typowy” znaczy tu bowiem zwykle tyle, co „najczęściej występujący”. Już z tego wynika, że istnieją przypadki inne niż typowe. Prawdziwość oceny, czy badany przypadek można porównywać z typowym, a tym bardziej szczegółowość i poprawność tej oceny, zależą od wielu czynników. Wśród nich niewątpliwie istotne są takie jak aktualny stan danej dziedziny wiedzy (wynikający z jej zaawansowania metodologicznego, liczby przeprowadzonych badań naukowych), liczba podobnych przypadków znanych biegłemu, liczby znanych mu przypadków nietypowych, wreszcie stan literatury naukowej (jej rozległości i jakości) z tej dziedziny, liczba i jakość ośrodków naukowych zajmujących się tą dziedziną. Znaczenie mają też stopień znajomości tego stanu wiedzy przez eksperta (czyli faktyczny zakres jego „wiadomości specjalnych”), jego doświadczenie zawodowe, ale także jego warunki osobiste, zdolności, a na koniec jego motywacje, podlegające najrozmaitszym wpływom – zarówno wewnętrznym, jak i zewnętrznym, w tym także sugestiom z różnych źródeł. Te sugestie płynąć mogą od osób prowadzących śledztwo albo być konsekwencją znajomości wyników innych ekspertyz przeprowadzonych w tej sprawie, sugerujących, że wynik aktualnie realizowanego badania nie powinien być z nimi sprzeczny. Wszystkie te czynniki składają się na margines subiektywizmu eksperta. W dużej mierze od jego doświadczenia zależy, czy trafnie rozpozna „typowość przypadku”, która jest później punktem odniesienia przy ocenie konkretnego przypadku. Wydaje się, że powyższe dostatecznie wyjaśnia, dlaczego biegli, dysponujący identyczną wiedzą podręcznikową, opierający się na identycznych danych wydają czasem rozbieżne opinie¹⁸.

Piszący – bodając jako jedyny w polskiej literaturze – o subiektywizmie w badaniach kryminalistycznych Jarosław Moszczyński zauważa, że: „w badaniach kryminalistycznych istnieją obszary subiektywizmu rozumiane jako obszary nie w pełni sprecyzowane, w których biegły posiada większą lub mniejszą swobodę interpretacji, ocen oraz dokonywania rozstrzygnięć”¹⁹.

Nie wdając się w logiczny rozbiór powyższego zdania (o jakie sprecyzowanie jakich obszarów chodzi? jakich rozstrzygnięć ma dokonywać biegły?), przyjmijmy, że przez element subiektywizmu w ekspertyzie (z zakresu nauk sądowych) rozumiemy, występujący przy ocenie dokonywanej przez eksperta, margines swobody jego interpretacji, niepodlegający żadnym obiektywnym kryteriom.

Szerokość tego marginesu jest różna przy różnych rodzajach ekspertyz. Przy daktyloskopii jest on relatywnie niewielki, a na przykład w ekspertyzie pisma – bardzo szeroki²⁰. Zależy to zarówno od stopnia złożoności materii, jak i od zaawansowania metodologicznego danej dyscypliny identyfikacji (badań).

Ostatnio w kryminalistyce zauważalny jest nurt kwestionujący dotychczasowe pojęcie identyfikacji indywidualnej²¹. Faktycznie, zalecenie, by biegły we wnioskach opinii pisał „uważam, że jest tak a tak” zamiast, jak to jest dziś powszechnie przyjęte, „stwierdzam, że jest tak a tak”, ma głębokie uzasadnienie metodologiczne i świadomie lub nie, uwzględnia margines subiektywizmu eksperta.

¹⁸ Por. K. Jaegermann, Z. Marek, *Rozbieżności w opiniach sądowo-lekarskich*, „Archiwum Medycyny Sądowej i Kryminologii” 1979, t. 29, nr 4, s. 249; por. K. Jaegermann, *op. cit.*, s. 67.

¹⁹ J. Moszczyński, *Subiektywizm w badaniach kryminalistycznych*, Olsztyn 2011.

²⁰ Por. H.H. Harralson, *Developments in handwriting and signature identification in the digital age*, Oxford, 2013, s. 120 i n., por. także: S.N. Srihari, S. Cha, H. Arora, S. Lee, *Individuality of handwriting*, „Journal of Forensic Sciences” 2002, t. 47, nr 4, s. 1–17; S.J. Strach, *Probability conclusions in handwriting comparison*, „International Journal of Forensic Document Examiners” 1998, t. 4, nr 4, s. 313–317.

²¹ Por. J. Konieczny, *Zmiana paradygmatu czy kryzys kryminalistyki?*, „Państwo i Prawo” nr 1, 2012, s. 3–16; idem, *Przeciwko kryminalistycznej identyfikacji indywidualnej*, „Problemy Współczesnej Kryminalistyki” 2014, t. 17, s. 49–58.

Zakres subiektywizmu przy ewaluacji zapisów na wykresach w badaniach poligraficznych

W badaniach poligraficznych przedmiotem oceny są zapisy krzywych wykreslanych przez poligraf, z których każda przedstawia zmiany w zakresie jednej z funkcji fizjologicznych organizmu osoby badanej. Standardowy poligraf rejestruje zmiany w przebiegu czynności oddychania, w pracy układu krążenia (zmiany częstotliwości tętna, względne zmiany ciśnienia krwi, zmiany reakcji skórno-galwanicznej, czasem dodatkowo inne parametry, jak na przykład zmiany reakcji pletysmograficznej, która zresztą jest konsekwencją modyfikacji w pracy układu krążenia)²². Są one wskaźnikami zmian aktywności organizmu, na które składają się zarówno wywołane pytaniami testu zmiany emocjonalne, jak i zmiany będące efektem wysiłku intelektualnego, koniecznego przy kłamstwie, które nazywamy najczęściej, choć niezupełnie ściśle „fizjologicznymi korelatami emocji”²³.

Zadaniem biegłego jest odróżnienie i ocena, które zmiany na krzywych są symptomatyczne i zostały wywołane treścią pytania testowego, a które są efektem rozlicznych artefaktów, a także porównywanie reakcji na pytania krytyczne (relevantne, związane – ang. *relevant*) i kontrolne (porównawcze – ang. *comparison*) lub obojętne (niezwiązane – ang. *irrelevant*). Biegły dokonuje zatem oceny zaobserwowanej zmiany przebiegu krzywych wykreslanych przez poligraf (rzadziej – odczytuje wielkości zmiany w reakcjach, wyrażone liczbowo w określonych jednostkach miary, np. w omach dla GSR/EDA), uznaje związek tej zmiany z treścią pytania, a na koniec – porównuje ją do jakiegoś wzorca. Jak widać – mamy tu do czynienia z dwoma ważnymi dla oceny elementami: uznaniem czegoś (jakiegoś kształtu krzywej) za wzorzec i dokładnym porównaniem zaobserwowanej w konkretnym badaniu zmiany z tym wzorcem.

Wspomniany wyżej wzorzec wraz z rozwojem techniki badań poligraficznych ulegał zmianie. W pierwszych latach stosowania poligrafu za symptomatyczną uznawano każdą zmianę w przebiegu krzywej, z wyjątkiem takiej, której przyczyna wydawała się oczywista i była inna niż pytanie testowe – na przykład: poruszenie się badanego, jego kichnięcie, kaszel – które to fakty odnotowywano zresztą na taśmie poligrafu. Tak więc każda zmiana przebiegu krzywej, jeśli nie była spowodowana wspomnianym poruszeniem się, kichnięciem, kaszlem itp., uznawana była za symptomatyczną. Natomiast wysoką częstotliwość, a zwłaszcza powtarzalność takich oczywistych zaburzeń zapisu oceniano z reguły jako próbę świadomego utrudniania badania, a więc coś, co przy końcowej ocenie wiarygodności badanego nie jest bez znaczenia²⁴.

Z czasem uznano, że „symptomami kłamstwa” są niektóre tylko zmiany w przebiegu krzywych. Powstawały całe atlasy ilustrujące przykładowe reakcje uznane za typowe. Takim swoistym atlasem były na przykład liczne ilustracje w kolejnych wydaniach książki Johna Reida i Freda Inbaua²⁵. Takimi atlasami, opracowanymi w oparciu o polskie badania eksperymentalne, były w pewnym sensie publikacje z końca lat 70.²⁶ Podobne powstają i dziś²⁷. Obecnie różne techniki badań uznają za symptomatyczne tylko niektóre zmiany w przebiegu krzywych, odmawiając takiego waloru zmianom do niedawna i w innych technikach uznawanych za symptomatyczne²⁸.

Jak widać, uznanie, co ostatecznie jest wzorcem dla reakcji symptomatycznej – pod wpływem doświadczeń praktyki i eksperymentów naukowych – również podlega ewolucji. Należy pamiętać, że ekspert w każdym konkretnym badaniu uzyskane reakcje porównuje z wzorcem – widać więc, że ta ocena, zazwyczaj nieoparta na dodatkowych pomiarach, jest jedynie interpretacją eksperta. Z prak-

²² Por. *Kryminalistyka*, red. J. Widacki, wyd. 3, Warszawa 2016, s. 418.

²³ *Ibidem*, s. 419.

²⁴ Por. J. Widacki, *Analiza przesłanek diagnozowania w badaniach poligraficznych*, Katowice 1982, s. 81–84.

²⁵ J. Reid, F. Inbau, *Truth and deception. The polygraph ("lie-detector") technique*, Baltimore 1977.

²⁶ Por. J. Widacki, *Z rozważań nad istotą symptomów kłamstwa przy lie-detection test*, „Archiwum Medycyny Sądowej i Kryminologii” 1975, t. 25, nr 1; idem, *Wartość diagnostyczna badania poligraficznego i jej znaczenie kryminalistyczne*, Kraków 1977 (ryc. 7, 8, 9).

²⁷ Por. np. А.Ю. Молчанов, Н.А. Молчанова, *Атлас Полиграмм*, Ярославль [A.J. Mołczanow, N.A. Mołczanowa, *Atlas poligramów*, Jarosław], 2007.

²⁸ Por. J. Widacki, *Historia badań poligraficznych*, Kraków 2017, s. 126–127.

tyki wiadomo, że eksperci oceniający te same zapisy poligraficzne dokonywali niekiedy odmiennych, nieraz sprzecznych ocen. Fakt istnienia takich rozbieżności słusznie przypisywano m.in. subiektywizmowi ekspertów.

Począwszy od lat 40. XX w., metodą tzw. ślepej interpretacji (ang. *blind scoring*) ustalano empirycznie zakres rozbieżności ocen ekspertów oceniających te same zapisy poligraficzne. Jak opisuje w niepublikowanej pracy doktorskiej Fabian Rouke²⁹, dwóch ekspertów, niezależnie od siebie, otrzymało do oceny zapisy reakcji uzyskanych przez innych ekspertów w czasie eksperymentu laboratoryjnego (były to jedynie zapisy reakcji skórno-galwanicznej). Ich zadaniem było dokonanie oceny reakcji i wskazanie „sprawcy”. Jeden ekspert wskazał „sprawcę” trafnie w 91% przypadków, drugi – w 88%. Zatem odsetek trafnych wskazań ekspertów różnił się tylko o 3 punkty procentowe³⁰. W innym eksperymencie³¹, sześciu ekspertów dostało do niezależnej oceny taśmy z zapisami z 25 autentycznych spraw (których wynik był znany), wyselekcjonowanych z archiwum policji w Miami. Rozpiętość wyników była tu większa niż w eksperymencie Roukego: od 69 do 81% trafnych ocen reakcji. Zatem rozbieżność trafności ocen wśród sześciu ekspertów wynosiła 12 p.p. Jeszcze większą rozpiętość wyników wykazał eksperyment Josepha Kubisa³²: od 73 do 92 % trafnych wskazań, czyli rozbieżność wynosiła 19 p.p. W badaniach Franka Horvatha i Johna Reida³³ taka rozbieżność między trafnymi ocenami ekspertów wynosiła aż 27,5 p.p. W kolejnych badaniach Horvatha³⁴ rozbieżność trafnych ocen między 10 ekspertami oceniającymi ten sam materiał wynosiła aż 31 p.p. W innych badaniach amerykańskich (Stanley Slowik i Joseph Buckley³⁵, Douglas Wicklander i Fred Hunter³⁶) różnica ocen dotyczyła odpowiednio 12,8 i 25 p.p.

W jedynych jak dotąd polskich badaniach³⁷ tego typu, z końca lat 70. XX w., dwóch ekspertów o identycznym przeszkoleniu i podobnym doświadczeniu zawodowym oceniało zapisy poligraficzne z badań eksperymentalnych 80 osób. Jeden z ekspertów uzyskał 85% trafnych wskazań, drugi – 67,5%. Jak widać, rozbieżność ich wyników wynosiła 17,5 p.p. Tabela 1 przedstawia podsumowanie wyżej opisanych rozbieżności między ekspertami w ocenach reakcji.

Należy zwrócić uwagę, że wszystkie te badania wykonane były z zastosowaniem jakościowych metod interpretacji zapisów. Pomijając pierwsze z badań – wykonane tylko na zapisach psychogalwanometru (najłatwiejszych do oceny) – rozbieżności trafności interpretacji reakcji w różnych badaniach wynosiły od 12 nawet do 31 p.p. Te różnice trafności ocen różnych ekspertów oceniających ten sam materiał mogły mieć różne przyczyny. Można zasadnie podejrzewać, że jedną z nich był subiektywizm ekspertów, rozumiany tak, jak zdefiniowaliśmy go wyżej.

²⁹ F.L. Rouke, *Evaluation of the indices of deception in the psychogalvanic technique*, 1941 (niepublikowana praca doktorska, Fordham University), [cyt. za:] G.H. Barland, *Reliability of polygraph chart evaluations*, „Polygraph” 1972, t. 1, nr 4, s. 192.

³⁰ Tak liczona rozbieżność ocen jest przybliżona i może mieć znaczenie tylko orientacyjne. Pokazuje w istocie jedynie, że eksperci, oceniając ten sam zbiór zapisów, różnili się w swoich ostatecznych ocenach.

³¹ W.D. Holmes, *The degree of objectivity in chart interpretation*, [w:] *Academy Lectures on Lie-detection*, t. 2, Springfield 1957, s. 67–70.

³² J.F. Kubis, *Studies in lie-detection. Computer feasibility considerations*, Fordham University, New York, 1962, RADC-TR 62-205, Project No 5534, AF 30(602)-2270, prepared for Rome Air Development Center, Air Force Systems Command, USAF, Griffiss AFB, New York.

³³ F. Horvath, J. Reid, *The reliability of polygraph examiner diagnosis of truth and deception*, „Journal of Criminal Law, Criminology and Police Science”, t. 62, nr 2, 1971, s. 276–281.

³⁴ F. Horvath, *The accuracy and reliability of Police polygraphic (lie-detector) examiners' judgements of truth and Deception. The effect of selected variables*, 1974 (niepublikowana praca doktorska, Michigan State University).

³⁵ S. Slowik, J.P. Buckley, *Relative accuracy of polygraph examiner diagnosis of respiration, blood pressure and GSR recordings*, „Journal of Police Science and Administration” 1975, t. 3, nr 3, s. 305–309.

³⁶ D.E. Wicklander, F.L. Hunter, *The influence of auxiliary sources of information in polygraph diagnosis*, „Journal of Police Science and Administration” 1975, t. 3, nr 4, s. 405.

³⁷ J. Widacki, *Wartość diagnostyczna badania poligraficznego...*

Tabela 1. Rozpiętość odsetka trafnych opinii ekspertów w wybranych badaniach naukowych

Autorzy (rok publikacji)	Rozpiętość ocen trafnych (%)	Rozbieżność ocen między ekspertami (p.p.)
Rouke (1941)	88–91	3
Holmes (1957)	69–81	12
Horvath, Reid (1971)	70–97,5	27,5
Kubis (1973)	73–92	19
Horvath (1974)	-	31
Słowik & Buckley (1975)	-	12,8
Wicklander & Hunter (1975)	70–95	25
Widacki (1977)	67,5–85	17,5

Źródło: opracowanie własne Autorów.

Badań, które bezpośrednio określałyby, do jakiego stopnia element subiektywny miał wpływ na dokładność wyniku badania poligraficznego i do jakiego stopnia oceny poszczególnych poligraferów dokonane na podstawie tego samego materiału są zgodne, jest stosunkowo niewiele. Jeden z pierwszych takich eksperymentów zrealizowano na zlecenie Departamentu Obrony USA³⁸: 30 poligraferów w ślepej interpretacji oceniało po 90 zapisów (po 30 wykonanych technikami Backstera, Reida i POT), z których połowa dotyczyła badanych zakwalifikowanych do zbioru DI (nieszczerých – ang. *deceptive*), połowa – do NDI (prawdomównych – ang. *non-deceptive*). Współczynnik korelacji Kappa w zależności od stosowanej techniki wynosił od 0,15 do 0,51³⁹. Najwyższą zgodność stwierdzono przy badaniach wykonanych techniką Backstera.

Podobny był zrealizowany kilka lat później eksperyment japoński, przeprowadzony przez Akihiro Suzuki i współpracowników⁴⁰: 26 ekspertów oceniało zapisy z 30 spraw realizowanych przez trzech ekspertów. Zgodność ocen mierzona współczynnikiem rzetelności Spearmann-Browna wynosiła 0,798.

W roku 2003 opublikowano raport opracowany dwa lata wcześniej dla Department of Defence Polygraph Institute⁴¹. Na podstawie analizy literatury porównuje on m.in. zgodność ocen (diagnoz) w badaniach psychologicznych, medycznych i poligraficznych. W badaniach psychologicznych ocena zgodności dotyczyła diagnoz DSM-III i DSM-IV, a w diagnostyce medycznej – badań rentgenologicznych, ultrasonograficznych, tomograficznych i rezonansu magnetycznego, a zatem zaawansowanych technologicznie metod diagnostyki. Według raportu, mierzona współczynnikiem Kappa zgodność ocen poligraferów wynosiła 0,77 i była porównywalna ze zgodnością ocen psychologów (która wynosiła 0,79) i była znacznie wyższa niż zgodność diagnoz lekarzy (która wynosiła jedynie 0,56).

Wszystkie ukazane wyżej wyniki zdają się potwierdzać tezę, że w badaniach poligraficznych, choć zgodność ocen ekspertów jest relatywnie wysoka – to zwłaszcza, gdy zapisy ocenia się jedynie jakościowo⁴², margines subiektywizmu eksperta jest szeroki. Ten margines subiektywizmu z założenia znacznie ograniczyć powinna numeryczna interpretacja zapisu.

³⁸ P.J. Bresh, R.A. Brisentine, *The reliability of blind interpretation of polygraph record for lie-detection purposes (Report prepared for DoD)*, 1968; [cyt. za:] J. Widacki, *Wprowadzenie do problematyki badań poligraficznych*, Warszawa 1981, s. 135–136.

³⁹ *Ibidem*.

⁴⁰ A. Suzuki, S. Watanabe, K. Ohnishi, K. Matsuno, M. Arasuna, *Polygraph examiners' judgements in chart interpretation: Reliability of judgement*, „Kagaku Keisatsu Kenkyusho” (Police Science Report) 1973, t. 26, nr 1, s. 34 i n.

⁴¹ P.E. Crewson, *A comparative analysis of polygraph with other screening and diagnostic tools*, „Polygraph” 2003, t. 32, nr 2, s. 57–85.

⁴² Posługujący się dziś tą niewątpliwie przestarzałą już metodyką interpretacji zapisów nazywają ją najczęściej „globalną” lub „holistyczną”.

Pierwsze próby numerycznej (ilościowej) interpretacji zapisów podjęto na początku lat 60. XX w. W badaniach, które zakończyły się raportem z 1962 r., J. Kubis posłużył się 4-stopniową skalą oceny reakcji. Reakcji bardzo wyraźnej (ang. *very significant*) przypisywano 3 punkty. Wyraźnej (ang. *significant*) – 2 punkty. Słabej, „wątpliwej” (ang. *doubtfully significant*) – 1 punkt. Gdy reakcji nie stwierdzono (ang. *non-significant*) – stawiano 0 punktów⁴³.

Jak podaje James Matte⁴⁴, przy realizacji eksperymentu Kubisa współpracował Cleve Backster. W 1963 r. Backster dopracował swoją technikę badań poligraficznych⁴⁵, której integralną częścią była numeryczna ocena zapisów według 7-stopniowej skali (od +3 do -3). Wedle tej skali Backster zalecał oceniać różnice między reakcjami na pary pytań krytycznych. Ocena miała być dokonywana na każdej krzywej oddzielnie, po czym wyniki podlegały sumowaniu. Jeśli reakcja na pytanie krytyczne była silniejsza niż reakcja na odpowiadające mu pytanie kontrolne, oznaczana była znakiem minus. Wielkość reakcji oceniano podobnie jak u Kubisa: 3 – różnica bardzo wyraźna, 2 – różnica wyraźna, 1 – różnica niewielka, 0 – brak różnicy. W przeciwieństwie do Kubisa, Backster wielkość cyfrową przypisywał nie wprost wielkości reakcji, ale różnicy między reakcją na pytanie krytyczne i odpowiadające mu pytanie kontrolne.

Wedle Backstera oceny należało dokonywać na każdej krzywej oddzielnie. Jeśli więc aparat był 3-kanałowy (pneumo, GSR, kardio), a w teście były trzy pary pytań krytycznych i kontrolnych, maksymalnie wartość reakcji mogła wynieść ± 27 pkt ($3 \times 3 \times 3$). Ze względu na to, że w technice Backstera wykonuje się ten sam test trzy razy, maksymalna liczba punktów mogła wynosić ± 81 (27×3). Inaczej mówiąc, wyniki każdego pojedynczego testu mieszczą się na kontinuum od -27 do +27 punktów, a wyniki całego badania – od -81 do +81 punktów. Wedle założeń Backstera, im wynik bliższy pozycji skrajnej, tym jest pewniejszy. Backster przyjął⁴⁶, że wynik między -5 a +5 dla pojedynczego testu, a między -15 a +15 dla całego badania, uznać należy za nierozstrzygnięty (INC – ang. *inconclusive*). Wynik między -15 a -81 – pozwala zaliczyć badanego do grupy „nieszczerych” (DI), wynik między +15 a +81 – do grupy „szczerych” (NDI).

Przy metodzie numerycznej oceny zapisów, ostateczna diagnoza (wskazanie) jest prostą konsekwencją wyniku liczbowego. Ten ostatni jest efektem sumowania punktów uzyskanych przy ocenie wielkości reakcji (ściślej: przy ocenie różnic reakcji na pytania krytyczne i kontrolne). Ocenie eksperta podlega tylko wielkość różnicy między reakcjami na pytanie krytyczne i odpowiadające mu pytanie lub pytania kontrolne. Ocena ta jest sformalizowana i ograniczona do skali (np. w systemie Utah: dramatyczna różnica; duża różnica; niewielka, ale zauważalna różnica; brak różnicy).

Do najważniejszych skal ocen numerycznych należy zaliczyć⁴⁷:

7- i 3-pozycyjne skale wykorzystywane przy ewaluacji zapisów na poligramach z testów pytań kontrolnych, w tym:

- 7-pozycyjną Rządu Federalnego Stanów Zjednoczonych (-3, -2, -1, 0, +1, +2, +3),
- 7-pozycyjną Uniwersytetu w Utah (-3, -2, -1, 0, +1, +2, +3),
- 3-pozycyjną Rządu Federalnego Stanów Zjednoczonych (-1, 0, +1),
- 3-pozycyjną w Empirycznym Systemie Oceniania (ESS – ang. *Empirical Scoring System*), zawierającą ten sam przedział ocen jak w federalnej skali 3-pozycyjnej, z wyjątkiem parametru EDA (ang. *electrodermal activity* – aktywność elektrodermalna), gdzie przypisuje się wartości: -2, 0, +2.

⁴³ J.F. Kubis, *op. cit.*, s. 20–22. Warto przypomnieć, że Kubis wprowadził tę skalę dla ułatwienia oceny zapisów poligraficznych wykonanych w ramach eksperymentu nie przez profesjonalnych poligraferów, ale przez studentów psychologii.

⁴⁴ J.A. Matte, *Forensic psychophysiology. Using the polygraph: scientific truth verification – lie detection*, New York 1996, s. 46.

⁴⁵ Por. J. Widacki, *Historia badań...*, s. 129.

⁴⁶ Wydaje się, że przyjął to arbitralnie, na podstawie własnego doświadczenia, brak bowiem publikowanych prac, z których wynikałoby statystyczne uzasadnienie dla takich założeń.

⁴⁷ Por. *Współczesne standardy badań poligraficznych*, red. M. Gołaszewski, Warszawa 2013, s. 26–43.

Modyfikacja ta jest rezultatem badań naukowych, które wykazały, że dane z tego parametru stanowią 50-procentowy wkład w ostateczną wartość decyzyjną i najlepiej korelują z kryterium „winy”;

- skalę ocen testu ukrytych informacji (CIT – ang. *Concealed Information Test*), z wykorzystaniem systemu Lykken’a (0, 1, 2).

W skalach 3-pozycyjnych przeznaczonych dla testów pytań kontrolnych obowiązuje prosta zasada: „większe jest lepsze” (ang. *bigger is better*). Natomiast skale 7-pozycyjne są nieco bardziej skomplikowane i w zależności od relatywnej wielkości reakcji istnieje więcej wariantów oceny. Każdą z powyższych skal uznaje się za potwierdzoną naukowo, ale federalna skala 3-pozycyjna nie spełnia obowiązujących standardów praktyki największego zrzeszenia ekspertów za zakresu badań poligraficznych na świecie – American Polygraph Association (APA). Testy oceniane wyłącznie tą skalą dają wyższy odsetek wyników nierozstrzygniętych niż określony przez APA limit 20%. Federalna skala 3-pozycyjna jest dopuszczalna pod warunkiem wykorzystania dodatkowo skali 7-pozycyjnej przy braku rozstrzygnięcia po pierwszym liczeniu.

W teście ukrytych informacji, zwanym także testem wiedzy o czynie (GKT – ang. *guilty knowledge test*), oceny przypisuje się na innych zasadach niż w testach pytań kontrolnych. W tym przypadku ukryty w pojedynczej serii testu bodziec krytyczny – „klucz” (np. nazwa lub cecha przedmiotu, sposób wykonania jakiejś czynności, fotografia miejsca czy osoby związanej ze sprawą) porównywany jest z bodźcami neutralnymi – równie prawdopodobnymi z punktu widzenia osoby niewinnej. Dana seria testu (podtest) jest oceniana na 2 – jeśli najbardziej znaczące zmiany reakcji fizjologicznych badanego wystąpiły przy bodźcu krytycznym; 1 – gdy była to druga pod względem istotności reakcja; 0 – w pozostałych przypadkach. Metoda ewaluacji Lykken’a opiera się wyłącznie na danych z komponentu EDA. Wynik testu równy liczbie poddanych ewaluacji podtestów oznacza identyfikację pozytywną (RI – ang. *recognition indicated* – stwierdzono rozpoznanie). Natomiast rezultat mniejszy od liczby podtestów zakwalifikowanych do oceny świadczy o braku rozpoznania kluczowego bodźca przez osobę badaną (NRI – ang. *no recognition indicated*). Opinia może być ponadto poparta liczbowo – poprzez określenie prawdopodobieństwa nierozpoznania szczegółów zdarzenia przez badanego.

Współcześnie numeryczna (ilościowo-jakościowa) interpretacja zapisów na poligramach jest immanentną i konieczną częścią wszystkich w zasadzie stosowanych technik badań poligraficznych. Eksperci korzystają głównie z trzech potwierdzonych naukowo systemów numerycznej analizy danych testowych (TDA – ang. *test data analysis*), obejmujących określone kryteria diagnostyczne przy interpretacji zapisów, rodzaje ocen i sposób ich łączenia oraz reguły decyzyjne. Te systemy skrótowo nazywamy: „federalnym”, „Utah” oraz „ESS”. Wszystkie są modyfikacjami metody opracowanej przez Backstera, a różnice występujące pomiędzy są nieznaczące.

System Rządu Federalnego USA opierał się najpierw na 22 cechach diagnostycznych nauczanych przez Szkołę Policji Wojskowej (USAMPS – ang. *United States Army Military Police School*)⁴⁸. Tak duża liczba kryteriów utrzymywała się od lat 70. XX w. aż do 2006 r. Wówczas Instytut Badań Poligraficznych Departamentu Obrony USA (DoDPI – ang. *Department of Defense Polygraph Institute*) – biorąc pod uwagę podstawy naukowe oraz ideę uproszczenia systemu – zredukował liczbę cech diagnostycznych do ośmiu głównych i trzech pomocniczych.

W latach 70. XX w. zespół naukowców z Uniwersytetu Utah (Salt Lake City), pod kierunkiem prof. Davida Raskina, rozpoczął pracę nad udoskonaleniem technik pytań kontrolnych. Już wtedy wiadomo było, że ocena numeryczna umożliwia istotnie większą trafność niż inne metody analizy danych z testów poligraficznych. Uznano natomiast, że znane dotąd systemy były niedoskonałe. Niektóre reguły decyzyjne czy cechy przyjmowane za diagnostyczne nie miały dostatecznych podstaw naukowych. Postanowiono więc zmodyfikować system oceny Backstera, który opierał się na zbyt skomplikowanych zasadach i był niekorzystny dla osób prawdomównych⁴⁹. W efekcie opracowano podejście „Utah” do testów pytań porównawczych wraz z systemem numerycznej analizy danych testowych,

⁴⁸ Zob. R.S. Weaver, *The numerical evaluation of polygraph charts: Evolution and comparison of three major systems*, „Polygraph” 1980, t. 9, nr 2, s. 94–108.

⁴⁹ B.G. Bell, D.C. Raskin, Ch.R. Honts, J.C. Kircher, *The Utah Numerical Scoring System*, „Polygraph” 1999, t. 28, nr 1, s. 1–9.

uwzględniającym dziewięć przesłanek diagnozowania. Koncepcja została potwierdzona wieloma badaniami i recenzowanymi publikacjami naukowymi w ciągu kolejnych 40 lat.

Najnowszą metodą jest tzw. Empiryczny System Oceniania, po raz pierwszy opisany w 2008 r. przez Raymonda Nelsona, Marka Handlera i Donalda Krapohla⁵⁰. Autorzy wyszli z założenia, że należy opracować system, który zapewni jak największą zgodność między ekspertami i trafność decyzyjną, będzie nieskomplikowany, doskonale udokumentowany naukowo, możliwy do skomputeryzowania i powszechnego stosowania zarówno przez doświadczonych, jak i niedoświadczonych poligrafów. Liczbę cech diagnostycznych ograniczono do sześciu. Do zalet systemu należy możliwość wskazania przez biegłego znaczenia statystycznego konkretnego rezultatu testu⁵¹. Wyniki końcowe numerycznej ewaluacji poligrafów uzyskuje się poprzez odniesienie się do ustalonych progów decyzyjnych, w zależności od przyjętej tolerancji błędu (standardowo 5–10%), wymaganego poziomu statystycznej istotności i prawdopodobieństwa błędu na podstawie reprezentatywnych danych⁵².

Wspólnymi cechami diagnostycznymi dla wszystkich trzech głównych systemów analizy danych testowych są: tłumienie oddechu i podniesienie linii bazowej krzywej oddychania; zmiany w amplitudzie w kanale EDA; podniesienie linii bazowej w kanale kardio oraz redukcja amplitudy pulsu w kanale reakcji naczynioruchowych.

Reguły decyzyjne w poszczególnych systemach ewaluacji danych zależą od typu testu (wieloprogowy, jednoprogowy, wieloaspektowy) i jego formatu (np. ZCT – ang. *Zone Comparison Test*; DLST – ang. *Directed Lie Screening Test*). Zostały wyznaczone w sposób arbitralny (jak u Backstera) lub na podstawie badań empirycznych, które dostarczyły danych do analiz statystycznych (jak w przypadku systemu ESS).

Warto pamiętać, że zapisy na poligramach są wizualną reprezentacją danych rejestrowanych przez poligraf. W przeszłości, kiedy stosowano aparaty analogowe, biegły mógł polegać wyłącznie na własnej obserwacji, obciążonej potencjalnymi zniekształceniami związanymi ze zdolnością postrzegania ludzkiego oka i dość dużą przestrzenią do subiektywnej interpretacji. Miał możliwość skorzystania co najwyżej z podręcznego narzędzia pomiaru w postaci linijki. Ponadto, nie mógł proporcjonalnie zwiększyć ani zmniejszyć krzywych na raz zadrukowanej rolce z wykresami. Jeśli w trakcie testu niedostatecznie wyregulował czułość jednego z czujników, mógł nie dostrzec istotnej różnicy w porównywanych reakcjach organizmu badanego. Długość linii oddechu (jedną z cech diagnostycznych) mógł sobie jedynie wizualizować w wyobraźni. Współcześnie, większość badań przeprowadza się z wykorzystaniem poligrafów komputerowych, współpracujących z oprogramowaniem, które eliminuje opisane trudności. Graficzne przedstawienie reakcji można dostosowywać zarówno na bieżąco w trakcie testu, jak i po jego zakończeniu (odbywa się to oczywiście automatycznie, z zachowaniem odpowiednich proporcji zarejestrowanych danych źródłowych, bez manipulacji). Opcje specjalistycznych programów pozwalają m.in. na pomiar amplitudy reakcji (wyrażonej liczbowo lub na podstawie graficznego wykresu) czy długości linii krzywej. Powstały algorytmy (np. do ewaluacji kanału reakcji naczynioruchowych), które dokładnie wyliczają różnice między fragmentami analizowanych zapisów. Bywa to niekiedy pomocne, choć np. reguły systemu ESS, w duchu uproszczenia procedury, zalecają powstrzymanie się od stosowania jakichkolwiek dodatkowych narzędzi pomiaru.

⁵⁰ R. Nelson, D.J. Krapohl, M. Handler, *Brute-force comparison: A Monte Carlo Study of the Objective Scoring System version 3 (OSS-3) and human polygraph scorers*, „Polygraph” 2008, t. 37, nr 3, s. 185–215.

⁵¹ Wartość diagnostyczna testu poligraficznego jest rozumiana jako uzyskana empirycznie (w badaniach naukowych obejmujących potwierdzone rozwiązania spraw) średnia dokładność (trafność) danego testu – czyli liczba prawidłowych identyfikacji, z wyłączeniem wyników nierozstrzygających, skoro te ostatnie nie skutkują żadną decyzją. Natomiast znaczenie statystyczne wyniku testu wylicza się na podstawie danych normatywnych właściwych dla populacji osób odpowiadających szczerze i nieszczerze (ewentualnie rozpoznających i nierozpoznających ukryte w teście szczegóły związane z określonym zdarzeniem). W oparciu o dane normatywne dotyczące poszczególnych rezultatów testów oraz przyjmowaną tolerancję błędu (wartość α – wynoszącą zazwyczaj 0,05 przy kwalifikowaniu reakcji jako typowe dla osoby odpowiadającej nieszczerze) ustala się optymalne progi decyzyjne przy analizie numerycznej zapisów reakcji badanego.

⁵² M. Handler, R. Nelson, W. Goodson, M. Hicks, *Empirical Scoring System: A cross-cultural replication and extension study of manual scoring and decision policies*, „Polygraph” 2010, t. 39, nr 4, s. 202.

Dzięki komputeryzacji potencjalny subiektywizm w zakresie samej interpretacji zapisów został istotnie ograniczony. Co więcej – stworzono programy dokonujące całościowej automatycznej ewaluacji danych i o ile mają one swoje ograniczenia wynikające np. ze sztywnego pola oceny reakcji i nie mogą zastąpić eksperta przy identyfikacji artefaktów – o tyle w analizowanym tu kontekście gwarantują pełen obiektywizm i perfekcyjną niezawodność. Niektóre z programów (np. OSS-3⁵³ czy algorytm dołączony do oprogramowania systemu poligrafu CPS) dostarczają wyników na poziomie trafności podobnym do manualnych ocen ekspertów, a bywa nawet, że je przewyższają⁵⁴.

Wprowadzenie numerycznej oceny zapisów i liczbowych kryteriów zaliczenia badanego do jednej z trzech grup: DI, NDI lub INC, wyeliminowało praktycznie te dodatkowe przesłanki diagnozowania w badaniach poligraficznych, które polegały na szeroko pojętej i z natury rzeczy jedynie jakościowej ocenie zachowania badanego przed badaniem, w jego trakcie i po nim⁵⁵. Poprzednio do tych symptomów przywiązywano dużą wagę i uwzględniano je przy podejmowaniu decyzji o końcowym wyniku badania⁵⁶.

Część empiryczna

Hipotezy badawcze

Biorąc pod uwagę przedstawione rozważania teoretyczne, sformułowano następujące hipotezy badawcze:

1. Metody numeryczne pozostawiają mniejszy margines subiektywizmu niż metody jakościowe (globalne, holistyczne).
2. Metody o mniejszej skali są bardziej obiektywne niż te o skalach 7-pozycyjnych. Pozostawiają bowiem mniejszy margines na subiektywną ocenę eksperta.
3. Oceniający metodą ślepej interpretacji – dzięki wyeliminowaniu innych niż zapisy na wykresach czynników (świadomie lub nieświadomie branych pod uwagę przy wydawaniu opinii) – uzyskają większy odsetek trafnych identyfikacji niż eksperci przeprowadzający badania.

Możliwe jest również przyjęcie hipotezy odwrotnej – tzn. że to badający będą dokładniejsi w swoich ocenach niż niezależni recenzenci, którzy nie mieli styczności z osobą badaną i nie znali przedmiotu badania. Oznaczałoby to, że w trakcie ekspertyzy, poza poligramami, biegły kieruje się mimo wszystko jeszcze innymi, subiektywnie odbieranymi przesłankami (np. niewerbalnymi, albo opiniami innych biegłych czy własną oceną faktów poznanych z akt sprawy) i co ważne – odbywa się to z korzyścią dla dokładności diagnozowania.

Na tę drugą możliwość zdaje się wskazywać znaczna część spełniających wysokie standardy naukowe analiz dotyczących trafności wskazań przy badaniach poligraficznych z wykorzystaniem technik pytań kontrolnych w rzeczywistych sprawach. Tabela 2 przedstawia wyniki testów ocenianych przez niezależnych ekspertów („na ślepo”), natomiast tabela 3 – rezultaty uzyskane przez poligrafików, którzy sami przeprowadzali badania. Odsetek prawidłowych wskazań osób winnych był zbliżony dla obu grup, natomiast badający byli trafniejsi od „ślepo” oceniających zapisy w przypadku wskazywania osób niewinnych.

⁵³ Por. R. Nelson, D.J. Krapohl, M. Handler, *op. cit.*

⁵⁴ Szerzej na ten temat: National Research Council, *The polygraph and lie detection*, Washington DC, 2003, s. 298–322; A.B. Dollins, D.J. Krapohl, D.W. Dutton, *A comparison of computer programs designed to evaluate psychophysiological detection of deception examinations: Bakeoff 1*, „Polygraph” 2000, t. 29, nr 3, s. 237–257.

⁵⁵ Por. J. Widacki, *Analiza przesłanek diagnozowania...*

⁵⁶ Por. F. Horvath, *Verbal and nonverbal clues to truth and deception during polygraph examinations*, „Journal of Police Science and Administration” 1973, t. 1, nr 2, s. 138–152.

Tabela 2. Wyniki badań terenowych na temat dokładności testów CQT ocenianych przez niezależne osoby

Autorzy* (rok publikacji)	Winni (%)				Niewinni (%)			
	n	Prawidłowo	Błędnie	Nierozstrzygnięte	n	Prawidłowo	Błędnie	Nierozstrzygnięte
Honts (1996)	7	100	0	0	6	83	0	17
Honts i Raskin (1988)	12	92	0	8	13	62	15	23
Patrick i Iacono (1991)	52	92	2	6	37	30	24	46
Raskin i in. (1988)	37	73	0	27	26	61	8	31
Średnia	27	89	1	10	82	59	12	29
Procent opinii		98	2	-		83	17	-

* Ch.R. Honts, *Criterion development and validity of the control question test in field application*, „Journal of General Psychology” 1996, t. 123, s. 309–324; Ch.R. Honts, D.C. Raskin, *A field study of the Directed Lie Control Question*, „Journal of Police Science and Administration” 1988, t. 16, s. 56–61; C.J. Patrick, W.G. Iacono, *Validity of the Control Question Polygraph Test: The problem of sampling bias*, „Journal of Applied Psychology” 1991, t. 76, nr 2, s. 229–238; D.C. Raskin i in., *A study of the validity of polygraph examinations in criminal investigations*, 1988 (raport końcowy dla National Institute of Justice, University of Utah).

Źródło: opracowanie własne na podstawie danych udostępnionych dzięki uprzejmości prof. D. Raskina.

Tabela 3. Wyniki badań terenowych na temat dokładności testów CQT ocenianych przez osoby przeprowadzające testy

Autorzy* (rok publikacji)	Winni (% prawidłowych identyfikacji)	Niewinni (% prawidłowych identyfikacji)
Horvath (1977)	100	100
Honts i Raskin (1988)	92	100
Raskin i in. (1988)	95	96
Patrick i Iacono (1991)	100	90
Honts (1996)	94	100
Średnia	98	97

* Jak w tab. 2 i dodatkowo: F. Horvath, *The effect of selected variables on interpretation of polygraph records*, „Journal of Applied Psychology” 1977, t. 62, nr 2, s. 127–136.

Źródło: opracowanie własne na podstawie danych udostępnionych dzięki uprzejmości prof. D. Raskina.

Opis eksperymentu

W finansowanym przez Narodowe Centrum Nauki projekcie naukowo-badawczym, zatytułowanym *Instrumentalne i nieinstrumentalne metody detekcji nieszczerości – problemy kryminalistyczne, etyczne i prawne*, realizowanym przez Krakowską Akademię im. Andrzeja Frycza Modrzewskiego, wzięło udział 15 ekspertów z zakresu badań poligraficznych. Trzech z nich przeprowadziło łącznie 39 badań osób, z których część odgrywała w eksperymencie rolę „winnych” (osoby, które oddały strzał na uniwersyteckiej strzelnicy sportowej), a pozostali nie mieli bezpośredniego związku z zainscenizowanym zdarzeniem.

Dla wzmocnienia motywacji, każdy z badanych, bez względu na rolę, jaką odgrywał w eksperymencie („winni” czy „niewinni”), otrzymywał banknot 50-złotowy. Osoby, które strzelały na strzelnicy i miały zataić ten fakt przed badającym poligraferem, były instruowane, że jeśli poligrafer wskaże je jako kłamie, będą musiały zwrócić 50 zł. Jeśli poligrafer nie rozpozna, że w czasie badania kłamały, 50 zł stanie się ich własnością. Z kolei osoby, które nie strzelały na strzelnicy i miały odgrywać rolę osób „niewinnych” i prawdopodobnych, były instruowane, że jeśli poligrafer potwierdzi ich prawdziwość, 50 zł stanie się ich własnością. Jeśli zaś poligrafer błędnie wskaże je jako kłamie, będą musiały zwrócić banknot. Miało to w założeniu przybliżyć motywację uczestników eksperymentu

do motywacji osób badanych w autentycznych sprawach. Dla osoby kłamiącej błąd poligrafera jest w nich korzystny, dla osoby prawdomównej – niekorzystny. Badający wykorzystywali test jednoprotblemowy⁵⁷ w formacie *Utah Zone Comparison Test* (test porównania stref, opracowany przez Uniwersytet Utah), o średniej dokładności wynoszącej 90,2–93% i średnim odsetku wyników nierozstrzygniętych na poziomie 9,8–10,7% (w zależności od rodzaju zastosowanych pytań porównawczych i systemu analizy danych testowych)⁵⁸.

Kolejnych 12 ekspertów miało za zadanie dokonać ślepej interpretacji zapisów na poligramach każdego z tych 39 badań. „Ślepi recenzenci” zostali podzieleni na trzy 4-osobowe grupy, z których każda miała posługiwać się inną metodą analizy danych testowych (globalną i dwoma numerycznymi: ESS i Utah). Dodatkowo wykorzystano całkowicie obiektywną metodę ewaluacji w postaci skomputeryzowanego algorytmu (reguły Sentera i analiza prawdopodobieństwa Raskina w programie OSS-3).

Poszczególne 4-osobowe grupy starano się dobrać tak, aby znajdowały się w nich osoby o podobnym poziomie wykształcenia i doświadczenia zawodowego, choć z obiektywnych przyczyn (przede wszystkim ze względu na relatywnie niewielką populację poligraferów w Polsce) – ten postulat mógł być uwzględniony jedynie częściowo. Najbardziej doświadczeni dokonywali interpretacji metodą Utah, a najmniej doświadczeni – ESS, ale wcześniejsze eksperymenty na świecie potwierdziły, że system ESS daje w przypadku niedoświadczonych poligraferów podobne rezultaty jak u wykwalifikowanych ekspertów, bez istotnych statystycznie różnic⁵⁹.

Wyniki i ich omówienie

Tabela 4 przedstawia rozpiętość trafnych identyfikacji między grupami ekspertów posługujących się różnymi metodami analizy danych testowych. Obliczono również współczynnik alfa Krippendorffa (tabela 5), bezpośrednio odnoszący się do zgodności między oceniającymi, z czym z kolei wiąże się zjawisko subiektywizmu. Co nie dziwi – niemal perfekcyjną zgodność uzyskano między wynikami ewaluacji z zastosowaniem algorytmów komputerowych w ramach programu OSS-3. Spośród oceniających manualnie – najbardziej zgodni byli ze sobą eksperci posługujący się systemem numerycznej interpretacji ESS (w skali 3-pozycyjnej). Większe rozbieżności odnotowano przy bardziej skomplikowanym systemie Utah (w skali 7-pozycyjnej) oraz ocenie globalnej, nazywanej szydlerczo „metodą na oko”.

Dla porównania – zgodność decyzji przy systemie ESS (zastosowanym dla testu jednoprotblemowego – takiego jak w opisanym eksperymencie), która została zmierzona współczynnikiem Kappa Fleiss’a i przedstawiona w opracowaniu M. Handler’a, R. Nelson’a, Walta Goodson’a i Matta Hicks’a, wyniosła 0,84 (95% CI = 0,73, 0,95)⁶⁰. Z kolei zgodność szczegółowych ocen numerycznych, przypisywanych każdemu mierzonemu parametrowi, wyniosła 0,56⁶¹–0,61⁶² u niedoświadczonych poligraferów i 0,57⁶³–0,61⁶⁴ u doświadczonych.

⁵⁷ W teście jednoprotblemowym pytania relewantne (krytyczne) są względem siebie zależne znaczeniowo. Nie ma wówczas logicznej możliwości, by badany, kłamiąc w odpowiedzi na jedno pytanie, był równocześnie szczery przy drugim. Odwrotnie jest w teście wieloprotblemowym (przesiewowym), gdzie pytania relewantne są od siebie niezależne znaczeniowo, dotyczą różnych zdarzeń i czynów, a zatem badany może jednocześnie na część pytań odpowiadać zgodnie z prawdą, a przy pozostałych próbować wprowadzić w błąd. Z podobną sytuacją mamy do czynienia w przypadku testu wieloaspektowego – gdzie pytania relewantne dotyczą wprawdzie jednego problemu, jednakże różnych jego aspektów (np. poza dokonaniem czynu także planowania, pomocnictwa czy wiedzy o nim – a więc okoliczności, które nie muszą współwystępować).

⁵⁸ Procedura przeprowadzania tego testu została szczegółowo opisana m.in. w: M. Handler, R. Nelson, *Utah approach to Comparison Question Polygraph Testing*, „Polygraph” 2009, t. 38, nr 1, s. 15–30.

⁵⁹ Zob. B. Blalock, B. Cushman, R. Nelson, *A replication and validation study on an empirically based manual scoring system*, „Polygraph” 2009, t. 38, nr 4, s. 281–288.

⁶⁰ M. Handler, R. Nelson, W. Goodson, M. Hicks, *op. cit.*, s. 205.

⁶¹ B. Blalock, B. Cushman, R. Nelson, *op. cit.*

⁶² R. Nelson, D.J. Krapohl, M. Handler, *op. cit.*

⁶³ *Ibidem*.

⁶⁴ R. Nelson, B. Blalock, M. Oelrich, B. Cushman, *Reliability of the Empirical Scoring System with expert examiners*, „Polygraph” 2011, t. 40, nr 3, s. 134.

Tabela 4. Rozpiętość trafnych opinii ekspertów w poszczególnych grupach w eksperymencie realizowanym w Krakowskiej Akademii im. Andrzeja Frycza Modrzewskiego w latach 2014–2017.

Metoda ewaluacji poligrafów	Rozpiętość trafnych identyfikacji (%)	Rozbieżność rezultatów między ekspertami (p.p.)
ESS	73,5– 86	12,5
Utah	79–96	17
globalna / jakościowa	69–83	27,5
algorytmy OSS-3	75,7–76,3	0,6

Źródło: opracowanie własne Autorów.

Tabela 5. Współczynniki zgodności (rzetelności) dla poszczególnych grup poligraferów w eksperymencie realizowanym w Krakowskiej Akademii im. Andrzeja Frycza Modrzewskiego w latach 2014–2017

Metoda ewaluacji poligramów	alfa Krippendorpha [CI]
ESS	.57 [.47, .66]
Utah	.43 [.33, .54]
globalna / jakościowa	.43 [.32, .53]
algorytmy OSS-3	.82 [.58, 1.00]

Źródło: opracowanie własne Autorów.

Rozbieżności w przypisywaniu przez ekspertów ocen numerycznych poszczególnym reakcjom organizmu (zapisom na poligramach) i przy podejmowaniu decyzji co do opinii końcowych mogły wynikać m.in. z takich czynników jak:

- niedookreśloność niektórych cech diagnostycznych⁶⁵ oraz reguł decyzyjnych – ich powiązanie ze zdolnościami percepcyjnymi eksperta (np. przy zasadzie „większe jest lepsze” wiele zależy od tego, w którym momencie oceniający dostrzeże różnicę i co uzna za „wyraźną” różnicę);
- błędne zastosowanie kryteriów diagnostycznych przewidzianych dla danego systemu analizy danych testowych;
- proste błędy matematyczne przy obliczeniach⁶⁶;
- kierowanie się przez eksperta innymi względami, niezwiązanymi z obiektywnymi przesłankami diagnozowania (błąd potwierdzenia⁶⁷, efekt pierwszeństwa⁶⁸, efekt aureoli⁶⁹ itp.).

⁶⁵ Chodzi tu np. o kontrowersje dotyczące złożoności reakcji elektrodermalnych – kiedy mówimy o reakcji złożonej, a kiedy już o dwóch odrębnych reakcjach?

⁶⁶ To może dość zaskakujące, ale przy analizie rozbieżności w przypisywaniu szczegółowych ocen numerycznych, okazało się, że w niektórych przypadkach eksperci prawidłowo zakwalifikowali zapisy jako odzwierciedlające reakcje symptomatyczne i dostrzegli wyraźną różnicę między reakcjami na pytanie krytyczne i kontrolne, lecz pomyłka przy zupełnie nieskomplikowanych obliczeniach powodowała, że wynik testu był inny, niż gdyby ekspert prawidłowo podsumował swoje oceny częściowe.

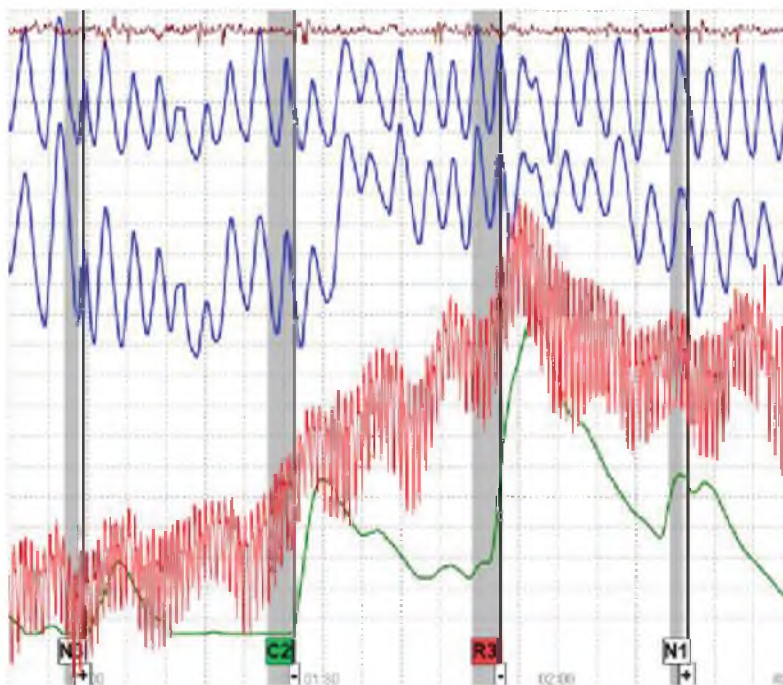
⁶⁷ Błąd potwierdzenia – preferowanie tych informacji, które potwierdzają wcześniejsze oczekiwania – niezależnie od tego, czy są prawdziwe.

⁶⁸ Efekt pierwszeństwa – większy wpływ informacji otrzymanych wcześniej (np. danych z akt, przypuszczeń śledczych) na tworzenie ogólnego wrażenia o osobie lub zdarzeniu, niż informacji, które przetworzyliśmy później.

⁶⁹ Efekt aureoli – tendencja do automatycznego przypisywania pozytywnych cech osobowości na podstawie pierwszego wrażenia.

Przykładowy fragment poligramu z zapisami niespójnie ocenianymi przez ekspertów przedstawia rysunek 1. Poligraferzy nie mieli w tym wypadku problemów z komponentem EDA (pierwsza krzywa od dołu). W przypadku kardio część uznała, że bardziej znacząca zmiana wystąpiła przy pytaniu krytycznym (R3) i przyznała ocenę -1, zaś pozostali ocenili widniejącą parę pytań neutralnie – tj. na 0⁷⁰. Natomiast zmiany cyklu oddechowego w strefie pytania R3 – w porównaniu do pytania kontrolnego C2 – były czasem interpretowane jako równorzędne, w innych przypadkach jako mniej znaczące, a w pozostałych – nawet bardziej, co skutkowało pełną gamą ocen numerycznych parametru pneumo wśród poligraferów⁷¹. Nie rozstrzygając, którzy z ocenających mieli rację, a którzy nie – należy zauważyć, że tego typu sporne sytuacje od czasu do czasu występują i wówczas niezwykle ważne może okazać się doświadczenie eksperta, uwzględnienie specyficznych cech osobniczych badanego czy tendencji w jego reakcjach na przestrzeni całego wykresu.

Rysunek 1. Fragment poligramu przedstawiający zapisy reakcji dla pary pytań: R3 i C2. Zapisy w kanałach pneumo i kardio były często odmiennie oceniane przez ekspertów



Źródło: opracowanie własne Autorów.

W tabeli 6 zestawiono odsetki trafnych identyfikacji oraz wyników nierozstrzygniętych w opiniach wydanych w oparciu o określoną metodę analizy zapisów na poligramach. Najlepiej poradzili sobie „ślepo” ocenający metodą numeryczną Utah (ok. 85%), a następnie metodą ESS (79%). Zbliżony poziom dokładności osiągnęli „ślepo” ocenający eksperci posługujący się metodą globalną i algorytma-

⁷⁰ Poligraferzy decydujący się na ocenę -1 dla parametru kardio argumentowali, że istotna zmiana w ciśnieniu krwi, którą można zaobserwować w okolicy pytania C2, zaczyna się jeszcze przed wprowadzeniem tego bodźca i dalej nie widać nowej reakcji fazowej. Natomiast zmiana przy pytaniu R2 jest bardzo wyraźna i zaczyna się w odpowiednim czasie po zadaniu pytania testowego. Co ciekawe – podobnych obserwacji zapisów dokonali także eksperci, którzy przyznawali ocenę 0. Powoływali się jednak na to, że skoro znacząca zmiana reakcji fizjologicznych nastąpiła jeszcze przed pytaniem C2, to mogło nie być dostatecznych warunków do tego, by organizm istotnie zareagował po wprowadzeniu bodźca, i dlatego pytanie C2 nie powinno być wykorzystane jako miarodajne pytanie porównawcze.

⁷¹ Przyznający ocenę +1 dostrzegali zmianę linii bazowej – zwłaszcza dla brzuszego (dolnego) pneumo w C2. Oceniający na 0 uznawali, że nie zaistniały jakieś konkretne cechy diagnostyczne, albo wystąpiły przy obu porównywanych pytaniach. Argumentacja dla oceny -1 opierała się na stwierdzeniu krótszej linii oddychania w strefie pytania R3.

mi komputerowymi oraz ci, którzy osobiście przeprowadzili badania (ok. 76–77%). Wyniki okazały się nieco niższe od tych przedstawionych dla testu Utah ZCT w raporcie APA z 2011 r., podsumowującym metaanalizę technik badań poligraficznych⁷². Należy przyjąć, że była to wypadkowa koncepcji eksperymentu, sposobu jego realizacji, kwalifikacji, doświadczenia i jakości pracy zaangażowanych poligraferów.

Odsetek wyników nierozstrzygniętych u sędziów kompetentnych był w eksperymencie wyższy od średniej w przypadku wykorzystania metody ślepej interpretacji zapisów (14,7–19,9%), zaś niższy u badających i w analizie za pomocą algorytmów komputerowych (odpowiednio: 3,8% i 5,1%). Stało się tak być może dlatego, że analiza komputerowa opierała się na precyzyjnych danych liczbowych, natomiast badający – w odróżnieniu do „ślepo” oceniających – usilniej dążyli do rozstrzygnięcia i brali pod uwagę dodatkowe, subiektywne czynniki przy opiniowaniu. Są to jednak jedynie prawdopodobne hipotezy, które wymagałyby udowodnienia.

Tabela 6. Trafność identyfikacji dokonanych przez poszczególne grupy poligraferów w eksperymencie realizowanym w Krakowskiej Akademii im. Andrzeja Frycza Modrzewskiego w latach 2014–2017

Metoda ewaluacji poligrafów		N (sprawca / świadek)	Nierozstrzygnięte (%)	Trafność* (%)	PPV** [CI]	NPV*** [CI]
	Badający (ESS)	39 (16/23)	5,1%	75,7%	.67 [.34, .90]	.76 [.64, .85]
„ślepa interpretacja”	ESS		19,9%	79%	.78 [.47, .94]	.89 [.73, .97]
	Utah		19,2%	84,5%	0.82 [.54, .97]	.82 [.68, .89]
	globalna / jakościowa		14,7%	76,7%	.62 [.37, .80]	.81 [.70, .92]
	algorytmy OSS-3		3,8%	76%	.75 [.39, .95]	.78 [.67, .83]

* odsetek trafnych identyfikacji z wyłączeniem wyników nierozstrzygających; ** prawdopodobieństwo, że osoba, która uzyskała pozytywny wynik testu, rzeczywiście posiada diagnozowaną cechę; *** prawdopodobieństwo, że osoba, która uzyskała negatywny wynik testu, rzeczywiście nie posiada diagnozowanej cechy.

Źródło: opracowanie własne Autorów.

Mając na uwadze wyniki powyższych analiz statystycznych – empirycznie potwierdzono, że metoda numeryczna ESS pozostawia istotnie mniejszy margines subiektywizmu niż metoda globalna (jakościowa). Ponadto, metody o mniejszej skali (w tym ESS) są bardziej obiektywne (z mniejszą przestrzenią na subiektywną ocenę eksperta) niż te o skali 7-stopniowej (w tym Utah).

Natomiast jeżeli chodzi o porównanie odsetka prawidłowych identyfikacji między ekspertami przeprowadzającymi badania a osobami dokonującymi ślepej interpretacji metodami numerycznymi, okazało się, że w ramach eksperymentu to ci ostatni byli dokładniejsi (84,5% dla ślepo oceniających systemem Utah i 79% dla ślepo oceniających systemem ESS wobec 75,7% w przypadku badających). Wydaje się, że mogło to wynikać z dwóch okoliczności:

- a) badający oceniali „na poczekaniu”, a „ślepo” interpretujący pracowali bez presji czasu w warunkach biurowych i domowych;

⁷² American Polygraph Association, *Meta-analytic survey of criterion accuracy of validated techniques*, „Polygraph” 2011, t. 40, nr 4, s. 195–305.

- b) nie można też wykluczyć, że badający – mniej lub bardziej świadomie – sugerowali się innymi (poza poligramami) czynnikami związanymi z osobistym odbiorem badanych, lecz wbrew dotychczasowym badaniom amerykańskim⁷³ – przyczyniło się to do obniżenia trafności identyfikacji.

Należy jednak zaznaczyć, że różnice w trafności analiz nie były na tyle statystycznie istotne, aby można było mówić o wyższości jednej metody nad innymi.

Generalnie w eksperymencie lepiej identyfikowano „niewinnych” niż „winnych”. Stało się tak prawdopodobnie m.in. dlatego, że pytania kontrolne odnosiły się do realnego życia, a krytyczne – jedynie do zainscenizowanego zdarzenia, ale w warunkach laboratoryjnych wydaje się to nieuniknione. Wykorzystana motywacja finansowa dla „winnego” badanego, za jego nietrafną identyfikację, mogła okazać się niewystarczająca, by zrównoważyć czy przeważać ciężar gatunkowy pytań związanych z realnymi przeżyciami, a wykorzystanych w testach jako kontrolne. Problem jest znany i opisywany w literaturze. Powszechnie przyjmuje się, że trafność wskazań w badaniach eksperymentalnych jest niższa niż w badaniach podejmowanych w realnych sprawach, właśnie ze względu na niższą motywację badanych⁷⁴. Zastosowanie wyższej motywacji badanych w ramach eksperymentu mogłoby się jednak okazać zbyt ryzykowne i niewspółmierne do efektów. Z kolei w warunkach dobrowolnego udziału w eksperymencie nie było możliwe zastosowanie sankcji negatywnej dla „zdemaskowanego kłamcy”. Z tego względu, że w badaniach eksperymentalnych pytania krytyczne dotyczą sytuacji wyreżyserowanej (umowne kłamstwo), fikcyjnej, a pytania kontrolne – prawdziwych zdarzeń z życia, tak jak w badaniach w realnych sprawach, emocjonalna waga tych ostatnich jest dla badanego z założenia większa. Dlatego też niektórzy autorzy sugerują, by w badaniach eksperymentalnych nie stosować testów pytań kontrolnych, a jedynie testy szczytowego napięcia (POT – ang. *peak of tension*) lub testy klasyczne⁷⁵. Jednak wielu autorów w badaniach eksperymentalnych z powodzeniem stosowało techniki pytań kontrolnych. Na przykład Gordon Barland w swych eksperymentach wykorzystywał technikę Backstera⁷⁶.

Wnioski

Badania, jak już napisano wyżej, wykazały, że numeryczne metody oceny zapisów poligraficznych są dokładniejsze i zmniejszają zakres subiektywizmu ekspertów. Wśród sposobów na ograniczenie subiektywizmu ekspertów z zakresu badań poligraficznych można wymienić standaryzację możliwie największej części, o ile nie całej procedury przeprowadzania badań. W tym kontekście komplementarne znaczenie ma komputeryzacja (automatyzacja instrukcji przekazywanych badanemu oraz zastosowanie algorytmów do analizy danych testowych).

Perspektywy są obiecujące, ale zalecane jest zachowanie ostrożności. Trzeba pamiętać, że badanie poligraficzne składa się z kilku integralnych części. Oprócz rejestracji reakcji i oceny uzyskanych zapisów, do zadań eksperta należy też warunkujące powodzenie odpowiednio sformułowanie pytań testowych, dobór właściwych testów, sposobu przeprowadzenia wywiadu przedtestowego (może z wyjątkiem standardowych elementów – takich jak zwięzły opis tego, co mierzy poligraf, wyjaśnienie najważniejszych kwestii związanych z psychologią i fizjologią człowieka) i rozmowy po testach. Tak więc samo perfekcyjne zbudowanie algorytmów ewaluacji zapisów poligraficznych, choć szalenie ważne i niewątpliwie ograniczające margines subiektywizmu, nie rozwiązuje problemu do końca. Badanie poligraficzne nie jest i zapewne nigdy nie będzie prostym pomiarem. Zawsze istotna będzie w nim rola eksperta – jego umiejętności, doświadczenie i predyspozycje.

⁷³ Por. np. W.D. Holmes, *op. cit.*

⁷⁴ Por. np. S. Abrams, *Polygraph validity and reliability. A review*, „Journal of Forensic Sciences” 1973, t. 18, nr 4, s. 318; por. także, J. Widacki, *Wprowadzenie do problematyki badań...*, s. 128–130.

⁷⁵ Por. F.K. Berrien, *A note on laboratory studies of deception*, „Journal of Experimental Psychology” 1939, t. 24, nr 5, s. 542–546. P.V. Trovillo, *Scientific proof of credibility*, „Tennessee Law Review” 1953, t. 22, s. 743–766.

⁷⁶ G.H. Barland, *An experimental study of field techniques in lie-detection*, 1972 (tekst niepublikowany, University of Utah); idem, *Detection of deception in criminal suspects. A field validation study*, 1975 (tekst niepublikowany, University of Utah).

Abstract

Expert's subjectivity in polygraph examination

The element of subjectivity in polygraph examination shall mean a margin of discretion, appearing in evaluation of expert's interpretation, which is not subordinated to any other objective criteria. The existence of this margin of discretion confirms a divergence of evaluations carried out by different experts assessing the same reactions. In assessing the same records of reactions, experts differed in their evaluation but the divergence of correct evaluations of the same reactions was ranging from 3% to 31%.

The convergence of expert's evaluations can be measured according to the reliability coefficient (e.g. Spearman-Brown), the correlation coefficient (e.g. Kappa, Fleiss' Kappa). The convergence of different expert's evaluations which is shown in literature (calculated according to Kappa coefficient) is relatively high (0.77).

In the experiment which was conducted by authors, experts was evaluating records of reactions using a qualitative method (global) and a quantitative method (numerical). The consistency coefficient (reliability) calculated by the alfa-Krippendorff coefficient was significantly higher applying the quantitative methods (numerical) – 0.43 to 0.82, whereas in the qualitative evaluation it was only 0.43.

In view of the above, authors' hypothesis is confirmed by the fact that there is the smaller scale of subjectivity in evaluations of numerical polygraphic records than in the qualitative method (global).

Key words: subjectivity of polygraph examination, subjectivity of experts opinion, qualitative and quantitative methods of polygrams interpretation

Streszczenie

Subiektywizm w badaniach poligraficznych

Przez element subiektywizmu w badaniach poligraficznych rozumiemy występujący przy ocenie dokonywanej przez eksperta margines swobody interpretacji, niepodlegający żadnym obiektywnym kryteriom.

O istnieniu tego marginesu przekonuje rozbieżność ocen dokonywanych przez różnych ekspertów. Oceniając te same zapisy reakcji, eksperci różnili się w ich ocenie, przy czym rozbieżność poprawnych ocen tych samych reakcji wynosiła od 3 do 31%.

Zgodność ocen ekspertów może być mierzona współczynnikiem rzetelności (np. Spearmana-Browna), współczynnikiem korelacji (np. Kappa, Kappa Fleissa).

Wykazywana w literaturze zgodność ocen różnych ekspertów (liczona współczynnikiem Kappa) jest relatywnie wysoka (0,77).

W wykonanym przez autorów eksperymencie, eksperci oceniali zapisy reakcji metodą jakościową („globalną”) oraz metodami ilościowymi (numerycznymi). Współczynnik zgodności (rzetelności) liczony współczynnikiem alfa-Krippendorffa był zdecydowanie wyższy przy zastosowaniu metod ilościowych (numerycznych) – 0,43 do 0,82 niż przy ocenie jakościowej – jedynie 0,43.

Potwierdza to hipotezę autorów, że numeryczne oceny zapisów poligraficznych są obciążone mniejszym marginesem subiektywizmu niż metoda jakościowa („globalna”).

Słowa kluczowe: subiektywizm w badaniach poligraficznych, subiektywizm opinii biegłego, ilościowe i jakościowe metody interpretacji poligramów

Część IV

Nieinstrumentalne metody detekcji kłamstwa

1. Wprowadzenie

Nieinstrumentalne metody detekcji kłamstwa, choć najczęściej tak nienazywane, stosowane były w praktyce kryminalistycznej od dawna. Do reguł sztuki śledczej należało obserwowanie osoby przesłuchiwanej, jej reakcji, zachowania, sposobu wypowiadania się, jej ruchów wyrazowych (tak mimicznych, jak i pantomimicznych). Podobnie zalecane było obserwowanie domowników w czasie przeprowadzania przeszukania (rewizji) pomieszczeń.

Wyniki tych obserwacji nie miały oczywiście waloru dowodowego, jednak pomagały funkcjonariuszom policji czy prokuratorom w zdobywaniu informacji przydatnych w czasie śledztwa.

Wspomniane czynności nie miały sformalizowanych ani procedur prawnych, ani taktyczno-kryminalnych. Opierały się wyłącznie na doświadczeniu zawodowym i zdolności obserwacyjnych je osób przeprowadzających.

Od czasu, gdy nieinstrumentalne metody detekcji kłamstwa zostały rozwinięte, a ich procedury sformalizowane i ujednolicone, pozostają do rozważenia dwie podstawowe kwestie. Pierwsza z nich, to jaki ewentualnie walor dowodowy mogą mieć informacje uzyskane w tej drodze. Kwestia druga to możliwość wprowadzenia tych metod do polskiego procesu karnego.

Odpowiedź na pierwsze z powyższych pytań wymaga przede wszystkim ustalenia wartości diagnostycznej wspomnianych metod. Drugie zagadnienie wymaga rozstrzygnięcia kilku kwestii prawnych.

2. Aktualny stan wiedzy

Przy opisie typów zeznań bardzo często stosowana jest kategoria podziału wprowadzona przez Simona de Laplace'a, na zeznania szczerze/nieszczere i prawdziwe/nieprawdziwe¹. Szczerłość to termin, który odnosi się do przekonania świadka co do tego, co i jak się wydarzyło, czyli zgodności treści wypowiedzi z reprezentacją pamięciową świadka. Prawdziwość natomiast to termin wskazujący na zgodność zeznań z prawdą obiektywną – z tym, co się tak naprawdę wydarzyło. Innymi słowy, zeznania świadka mogą być szczerze (świadek zeznaje

¹ B.W. Wojciechowski, *Ocena wiarygodności zeznań świadków dorosłych przy zastosowaniu psychologicznych i prawnych kryteriów treściowych*, niepublikowana rozprawa doktorska, Uniwersytet Śląski w Katowicach, 2012.

to, co pamięta i jest przekonany, że pamięta zdarzenie dokładnie), ale nieprawdziwe (niezgodne z prawdą – bo na przykład świadek pewne elementy zdarzenia źle lub niedokładnie pamięta), lub stanowić dowolną kombinację tych cech. W przypadku metod nieinstrumentalnych możliwe jest wskazywanie kłamstwa w rozumieniu „nieszczerości” – czyli sytuacji, kiedy w sposób świadomy dana osoba podaje informacje, co do których jest przekonana, że są nieprawdziwe.

Od wieków ludzie są przekonani, że zjawisku kłamania towarzyszą szczególne, obserwowalne zachowania, dzięki którym można je zidentyfikować. Co więcej, to przekonanie jest również popularne wśród naukowców. Na przykład Zuckerman, DePaulo i Rosenthal zauważyli, że w zachowaniu werbalnym i niewerbalnym mogą pojawiać się tzw. wskaźniki kłamstwa, które są związane z przeżywaniem przez kłamców szczególnych stanów: (1) ogólnego pobudzenia, (2) szczególnych rodzajów emocji doświadczanych podczas kłamania, (3) przeciążenia procesów poznawczych oraz (4) próby kontrolowania zachowania podczas kłamania².

Po pierwsze, kłamcy w momencie kłamania mogą przeżywać większe niespecyficzne pobudzenie w porównaniu do osób mówiących prawdę. Jest ono związane z aktywnością układu nerwowego, który stawiany jest w stan gotowości w momencie kłamania. Może się to przejawiać na przykład w przyspieszeniu tempa wypowiedzanych słów, zwiększeniu potliwości skóry, szybszym biciu serca i wyższym tonie głosu. Po drugie, kłamcy mogą przeżywać większe natężenie szczególnych emocji: strachu (przed złapaniem na kłamstwie), podekscytowania lub poczucia winy (z powodu możliwości wprowadzenia kogoś w błąd). Z tego względu mogą przejawiać zachowania unikowe i wskazujące na stres, np. drżenie dłoni czy zaczerwienienie skóry. Po trzecie, kłamcy doświadczają dużo większego obciążenia poznawczego niż osoby szczerze. Kłamanie jest bardzo skomplikowaną czynnością: kłamca musi pamiętać, co mówi, i nie zaprzeczać samemu sobie. Musi przewidzieć, co jego rozmówca może wiedzieć o sprawie i dopasowywać do tego swoje wypowiedzi. Dodatkowo musi stale aktywnie hamować prawdziwe wspomnienia dotyczące danej sprawy i wymyślać na bieżąco nowe, fałszywe szczegóły. Wszystko to sprawia, że kłamanie jest bardzo trudnym poznawczo zadaniem i może przejawiać się np. w dłuższym czasie potrzebnym kłamcy na odpowiedź na pytanie, zubożeniem gestykulacji itp. Po czwarte wreszcie kłamcy zdają sobie sprawę z tego, że podczas kłamania mogą występować w ich zachowaniu różne wskazówki kłamania, dlatego starają się nad nimi panować. Starają się kontrolować zachowanie, tłumić gesty, które uważają za zdradzające kłamstwo – może się to przejawiać w sztywnej postawie ciała lub obniżonej ekspresji emocjonalnej.

Na podstawie tej czteroczynnikowej teorii przyjęto, że mogą istnieć różne wskaźniki kłamstwa, zarówno obecne w fizjologii (nadmierna potliwość skóry może być związana ze wzrostem ogólnego pobudzenia), w zachowaniu niewerbalnym (np. ogólna nieruchomość dłoni może być związana z nadmierną kontrolą swojego zachowania, a co za tym idzie – z kłamaniem) oraz werbalnym (np. mała liczba szczegółów jest związana z obciążeniem poznawczym, czyli brak szczegółów w wypowiedzi może świadczyć o kłamstwie). Analogicznie, zaczęto systematyzować znane metody wykrywania kłamstwa ze względu na analizowany kanał komunikacji³: metody oparte na analizie psychofizjologii (poligraf, badania EEG czy fMRI, badania termograficzne), metody oparte na analizie niewerbalnego oraz parawerbalnego zachowania⁴ oraz metody polegające na analizie werbalnych wskazówek, czyli głównie

² M. Zuckerman, B.M. DePaulo, R. Rosenthal, *Verbal and nonverbal communication of deception*, [w:] *Advances in experimental social psychology*, Vol. 14, ed. L. Berkowitz, Academic Press, New York 1981, s. 1–60.

³ A. Vrij, *Detecting lies and deceit: pitfalls and opportunities*, Wiley, Chichester 2008.

⁴ B.M. DePaulo, J.J. Lindsay, B.E. Malone, L. Muhlenbruck, K. Charlton, H. Cooper, *Cues to deception*, „Psychological Bulletin” 2003, Vol. 129, s. 74–118.

treści zeznania⁵. Dwa ostatnie typy są czasem określane jako metody nieinstrumentalne⁶ ze względu na fakt, że do ich stosowania nie jest konieczna specjalistyczna aparatura.

2.1. Obserwacja zachowania niewerbalnego

Współcześnie raczej niewiele badań jest prowadzonych nad niewerbalnymi i parawerbalnymi (wokalnymi) wskaźnikami kłamstwa. Po kilkudziesięciu latach badań i eksperymentów mniej więcej wiadomo, że nie ma typowego i stałego zachowania niewerbalnego, które towarzyszyłoby kłamaniu. Dodatkowo trafność identyfikacji kłamstwa w oparciu o takie wskazówki jest bardzo niska⁷. Dlaczego tak się dzieje?

Głównie dlatego, że ludzie na ogół zwracają uwagę na niewłaściwe wskazówki. Popularność różnych poradników, szkoleń i stron internetowych, na których można nauczyć się jak „skutecznie” wykrywać kłamstwo, przyczyniła się do powstania wielu mitów i rozpowszechnienia nieprawdziwej wiedzy. W ramach badań The Global Deception Research Team⁸ przeanalizowano przekonania dotyczące najbardziej rozpowszechnionych wskazówek kłamania wśród 2320 osób z 58 krajów. Większość z podawanych przez badanych jako „najbardziej trafne” nie pokrywała się z tymi, które zostały zidentyfikowane w badaniach naukowych. Na przykład 64% badanych wskazywało unikanie wzroku jako trafną wskazówkę kłamania – wskazówkę, która już od dawna jest określana w badaniach naukowych jako nietrafna⁹. Dodatkowo badacze ci, dokonując przeglądu badań, zauważyli, że ludzie powszechnie uważają unikanie wzroku, ruchy ciała, automanipulacje oraz drżenie rąk za skuteczne wskazówki kłamstwa. Tymczasem metaanaliza DePaulo i współpracowników wykazała, że trafność tych wskazówek jest niemal równa zeru (odpowiednio $d = 0,01$, $d = 0,05$, $d = -0,01$ oraz $d = -0,12$)¹⁰. Co więcej, podobnie błędne przekonania podzielają specjaliści, tacy jak policjanci czy ławnicy¹¹.

Aby wskazać użyteczność narzędzi, zwyczajowo w badaniach naukowych mierzy się dwa parametry: wielkość efektu (trafność) dla każdej wskazówki kłamstwa oraz trafność ocen dokonywanych przez ludzi wyszkolonych w używaniu danej wskazówki. Wielkość efektu dla danej wskazówki kłamstwa to po prostu siła, z jaką jest związana z kłamaniem/szczerością.

⁵ Np. Content Based Criteria Analysis – CBCA, M. Steller, G. Kohnken, *Statement analysis: Credibility assessment of children's testimonies in sexual abuse cases*, [w:] *Psychological methods in criminal investigation and evidence*, eds. D.C. Raskin, Springer, New York 1989, s. 217; Reality Monitoring – RM, S.L. Sporer, *Reality monitoring and the detection of deception*, [w:] *The detection of deception in forensic contexts*, eds. P.-A. Granhag, L.A. Strömwall, Cambridge University Press, Cambridge 2004, s. 64–102; Aberdeen Report Judgment Scale – ARJS, S.L. Sporer, M.M. Breuer, *Validity of the Aberdeen Report Judgment Scales-Short Training Version-German (ARJS-STV-G)*, Paper presented at the annual meeting of the American Psychology – Law Society, TBA, San Antonio 2005.

⁶ Np.: J. Widacki, *Historia badań poligraficznych*, Oficyna Wydawnicza Krakowskiej Akademii im. Andrzeja Frycza Modrzewskiego, Kraków 2017.

⁷ C.F. Bond Jr., B.M. DePaulo, *Accuracy of deception judgments*, „Personality and Social Psychology Review” 2006, Vol. 10, s. 214–234.

⁸ Global Deception Research Team, *A world of lies*, „Journal of Cross-Cultural Psychology” 2006, Vol. 37, s. 60–74.

⁹ C.F. Bond, T.R. Levine, M. Hartwig, *New findings in non-verbal lie detection*, [w:] *Detecting deception*, eds. P.A. Granhag, A. Vrij, B. Verschuere, Wiley Blackwell, Chichester 2015, s. 38–58.

¹⁰ B.M. DePaulo, J.J. Lindsay, B.E. Malone, L. Muhlenbruck, K. Charlton, H. Cooper, *Cues to deception...*, *op. cit.*

¹¹ Np.: L.H. Colwell, H.A. Miller, R.S. Miller, P.M. Lyons, *US police officers' knowledge regarding behaviors indicative of deception: implications for eradicating erroneous beliefs through training*, „Psychology, Crime & Law” 2006, Vol. 12, s. 489–503; J. Widacki, N. Mirska, M. Wrońska, *Werbalne i niewerbalne symptomy kłamstwa w oczach policjantów oraz psychologów*, „Przegląd Bezpieczeństwa Wewnętrznego” 2012, nr 7, s. 19–30; J. Ulatowska, *Deklarowana wiedza „ekspertów” i „laików” na temat przejawów kłamstwa*, „Psychologia Jakości Życia” 2005, nr 4(2), s. 165–180; J. Ulatowska, *Knowledge of cues to deception – looking for its determinants*, „Problems of Forensic Sciences” 2009, Vol. 80, s. 411–428.

W metaanalizach zwykle tą siłę podaje się za pomocą parametru d Cohena – jeżeli parametr ten jest poniżej 0,10 uznaje się, że efektu nie ma, przy 0,30 efekt uznawany jest za średni a przy 0,50 za silny¹².

Zwykle to jednak nie wystarcza, by powiedzieć, czy wskazówka jest pomocna w ocenianiu szczerości zeznań. Można sobie bowiem wyobrazić sytuację, kiedy wskazówka jest bardzo mocno związana z kłamaniem, ale osoba dokonująca analizy zeznania i tak ostatecznie dokonuje błędnej decyzji co do jego szczerości. Może się tak dzieć np. na skutek błędów w procesach decyzyjnych, na skutek tego, że wskazówka jest trudna do analizy itd. Dlatego oprócz wielkości efektu dla każdej wskazówki wylicza się ogólną trafność dla osób dokonujących oceny przy użyciu danej wskazówki.

Najpopularniejszym jej wskaźnikiem jest dokładność (zwana również po prostu trafnością) czyli określenie, ile zeznań szczerych i nieszczerych za pomocą narzędzia udało się poprawnie zaklasyfikować, w proporcji do sumy zeznań trafnie i nietrafnie sklasyfikowanych. Zwykle wylicza się również czułość (czyli jak wielu kłamców poprawnie zaklasyfikowano) i swoistość (ile osób szczerych poprawnie sklasyfikowano). Rozróżnienie to jest istotne, bo w praktyce często obserwuje się przetarg między czułością i swoistością – narzędzia (lub wskazówki), które na ogół zwiększają czułość jednocześnie zmniejszają swoistość. Innymi słowy, im więcej kłamców udaje się wykryć, tym więcej osób szczerych niesłusznie posądza się o kłamstwo.

Metaanalizy dotyczące niewerbalnych wskazówek kłamania¹³ wskazują na to, że większość z nich ma bardzo mały efekt. Najczęściej wykazywane w badaniach zachowania, takie jak długość wypowiedzi czy utrzymywanie kontaktu wzrokowego, są bardzo słabo powiązane z kłamaniem (odpowiednio $d = 0,02$ oraz $d = 0,01$), a z kolei te rzadziej badane, jak ruchy stóp, wykazują wprawdzie relatywnie dużą wielkość efektu ($d = 1,05$), ale z kolei są niezwykle rzadko spotykane. Co więcej, wiele z tych wskazówek jest niekonsekwentnych – w niektórych badaniach są wskazówkami kłamania, a w innych wskazówkami szczerości¹⁴.

Wydaje się, że dużo lepsze od niewerbalnych są tzw. parawerbalne wskazówki, takie jak wysokość głosu, czas latencji między pytaniem a odpowiedzią czy błędy w wypowiedzi. Metaanaliza Sporer'a i Schwandt¹⁵ wskazuje na to, że przynajmniej kilka z tych wskazówek jest w umiarkowanie silny sposób faktycznie związanych z kłamstwem. Niestety, bazowanie na nich przy próbie identyfikacji kłamcy nie prowadzi do zwiększenia trafności, czułości i swoistości, które zwykle nie przekraczają poziomu przypadku¹⁶. Co więcej, metaanaliza badań nad skutecznością szkoleń poświęconych wykrywaniu kłamstwa w oparciu o wskazówki niewerbalne i parawerbalne pokazuje, że nawet intensywne szkolenia nie poprawiają znacząco skuteczności osób biorących w nich udział¹⁷.

Reasumując wszystkie metaanalizy i badania, można powiedzieć, że korelacje pomiędzy niewerbalnymi i parawerbalnymi wskazówkami kłamstwa a kłamstwem są bardzo niskie lub raczej niskie i zależą od bardzo dużej liczby moderatorów, takich jak motywacja kłamcy, treść zeznań, sposób przygotowania do przesłuchania itd.¹⁸. Dodatkowo są bardzo nietrafne są

¹² J. Cohen, *Statistical power analysis for the behavioral sciences*, 2nd ed., Lawrence Erlbaum, Hillsdale 1988.

¹³ B.M. DePaulo, J.J. Lindsay, B.E. Malone, L. Muhlenbruck, K. Charlton, H. Cooper, *Cues to deception...*, *op. cit.*

¹⁴ C.F. Bond, T.R. Levine, M. Hartwig, *New findings...*, *op. cit.*

¹⁵ S.L. Sporer, B. Schwandt, *Paraverbal indicators of deception: A meta-analytic synthesis*, „Applied Cognitive Psychology” 2006, Vol. 20, s. 421–446.

¹⁶ C.F. Bond Jr., B.M. DePaulo, *Accuracy of deception...*, *op. cit.*

¹⁷ V. Hauch, S.L. Sporer, S.W. Michael, C.A. Meissner, *Does training improve detection of deception? A meta-analysis*, „Communication Research” 2016, Vol. 43, s. 283–343.

¹⁸ B.M. DePaulo, J.J. Lindsay, B.E. Malone, L. Muhlenbruck, K. Charlton, H. Cooper, *Cues to deception...*, *op. cit.*; S.L. Sporer, B. Schwandt, *Paraverbal indicators...*, *op. cit.*; S.L. Sporer, B. Schwandt, *Moderators of non-verbal indicators of deception: A meta-analytic synthesis*, „Psychology, Public Policy, and Law” 2007, Vol. 13, s. 1–34.

próby wykrycia kłamstwa na podstawie obserwacji zachowania niewerbalnego¹⁹, a intensywne szkolenia poświęcone wykrywaniu kłamstwa w oparciu o wskazówki niewerbalne niemal nie zwiększają trafności²⁰. Jest to o tyle ciekawe, że w powszechnej świadomości, również specjalistów, niewerbalne wskazówki kłamstwa są najbardziej skuteczne. Jednakże brak empirycznych podstaw ich stosowania sprawił, że obecnie w psychologii najwięcej miejsca poświęca się usystematyzowanym, wystandaryzowanym narzędziom opartym na analizie werbalnych aspektów zeznania – konkretnie jego treści

2.2. Werbalne narzędzia oceniania szczerości zeznań

Już w latach pięćdziesiątych XX wieku zaczęły powstawać pierwsze systemy oceniania zeznań, których celem było wykrycie nieszczerości tylko na podstawie treści zeznania i elementów, które się w nim pojawiały. W ramach tych systemów zwykło się przyjmować, że dla oceny szczerości zeznania istotna jest tylko treść wypowiedzi, z pominięciem wskazówek parawerbalnych i niewerbalnych. Dlatego też analiz dokonuje się zwykle na spisanych dosłownych transkrypcjach zeznania świadka. Stąd też czasem metody te określane są w literaturze anglosaskiej jako werbalne narzędzia oceny wiarygodności/szczerości zeznań (*verbal veracity assesment tools*) lub treściowe metody oceny szczerości²¹.

Zwykle jako podstawę teoretyczną istnienia metod treściowych podaje się tzw. hipotezę Undeutch'a – założenie, że zeznania oparte na osobistych doświadczeniach świadka będą różniły się pod względem formy i treści (jakościowo i ilościowo) od zeznań nieszczerých oraz że motywacja kłamcy jest inna niż osoby mówiącej prawdę, co przekłada się na treść zeznania²². Na podstawie tej hipotezy stworzono kilka systemów oceny zeznań na podstawie jego treści oraz przynajmniej kilkadziesiąt kryteriów dotyczących cech pojawiających się częściej w zeznaniach szczerých niż nieszczerých²³. Prawdopodobnie najczęściej stosowanym w praktyce wymiaru sprawiedliwości systemem jest SVA – Analiza Szczerości Zeznań (*Statement Validity Assesment*) Stellera i Kohnkena²⁴. Procedura ta obejmuje 1) analizę akt sprawy, 2) przeprowadzenie ustrukturyzowanego przesłuchania, 3) analizę treści tego zeznania (czyli CBCA, Opartej na Kryteriach Analizie Treści, *Content Based Criteria Analysis*) oraz 4) sprawdzenie Listy Kontrolnej Prawdziwości (*Validity Checklist*).

CBCA jest jednym z najlepiej zbadanych narzędzi wykrywania kłamstwa bazujących na analizie treści zeznań. Dotychczas opublikowano na jego temat 4 metaanalizy²⁵ oraz ponad 80 artykułów naukowych (stan na 2017 rok²⁶). CBCA składa się z 19 kryteriów dotyczących semantycznych cech treści zeznania, które częściej pojawiają się w zeznaniach szczerých (dotyczących zdarzenia, którego osoba rzeczywiście była świadkiem) niż nieszczerých (dotyczą-

¹⁹ C.F. Bond Jr., B.M. DePaulo, *Accuracy of deception...*, *op. cit.*

²⁰ V. Hauch, S.L. Sporer, S.W. Michael, C.A. Meissner, *Does training improve detection...*, *op. cit.*

²¹ A. Vrij, *Criteria-based content analysis: A qualitative review of the first 37 studies*, „Psychology, Public Policy, and Law” 2005, Vol. 11, s. 3–41.

²² F. Arntzen, *Psychologia zeznań świadków*, Państwowe Wydawnictwo Naukowe, Warszawa 1989.

²³ Np.: *ibidem*; S.H. Adams, *Statement analysis: What do suspects' words really reveal?*, FBI Law Enforcement Bulletin, 1996; Z. Marten, *Psychologia zeznań*, Lexis Nexis, Warszawa 2012.

²⁴ M. Steller, G. Kohnken, *Statement analysis...*, *op. cit.*

²⁵ B.G. Amado, R. Arce, F. Fariña, *Undeutsch hypothesis and Criteria-Based Content Analysis: A meta-analytic review*, „The European Journal of Psychology Applied to Legal Context” 2015, Vol. 7, s. 3–12; B.G. Amado, R. Arce, F. Fariña, M. Vilariño, *Criteria-Based Content Analysis (CBCA) reality criteria in adults: A meta-analytic review*, „International Journal of Clinical and Health Psychology” 2016, Vol. 16, s. 201–210; V. Hauch, S.L. Sporer, J. Massip, I. Blandon-Gitlin, *Can credibility criteria be assessed reliably? A meta-analysis of criteria-based content analysis*, „Psychology Assessment” 2017, Vol. 29, s. 819–834; V. Oberlander, C. Naefgen, J. Koppehele-Gossel, L. Quinten, R. Banse, A.F. Schmidt, *Validity of content-based techniques to distinguish true and fabricated statements: A meta-analysis*, „Law and Human Behavior” 2016, Vol. 40, s. 440–457.

²⁶ Dane za: V. Hauch, S.L. Sporer, J. Massip, I. Blandon-Gitlin, *Can credibility criteria...*, *op. cit.*

cych zmyślonych zdarzeń). Te kryteria to np. logiczna struktura wypowiedzi (kryterium 1), duża liczba szczegółów (kryterium 3) czy opisy interakcji (kryterium 5). Pojawianie się tych cech w zeznaniu szczerym związane jest z funkcjonowaniem epizodycznej pamięci autobiograficznej u osób szczerych (podczas gdy osoby nieszczerze przy tworzeniu zeznania korzystają raczej ze skryptów i schematów) oraz ze stosowanymi przez osoby szczerze strategiami autoprezentacji²⁷.

Analiza szczerości zeznania za pomocą CBCA polega zwykle na tym, że przeszkolony specjalista poszukuje każdego z 19 kryteriów w danym zeznaniu, a następnie przypisuje im wartość punktową. Zwykle jest to 0 – jeżeli dane kryterium nie występuje, 1 – jeżeli występuje w słabym nasileniu, lub 2 – jeżeli występuje w sposób bardzo wyrazisty²⁸. Stosowana jest też skala pięciopunktowa²⁹ i siedmiopunktowa³⁰. Co ważne, nie powinno się sumować wszystkich punktów – ostateczna ocena co do szczerości zeznania podejmowana jest przez biegłego na podstawie wagi poszczególnych kryteriów oraz pewnych systematycznych reguł decyzyjnych. Np. Craig, Sheibe, Kircher, Raskin i Dodd uważają, że jeżeli pięć dowolnych kryteriów jest obecnych, można uznać, że zeznanie jest prawdziwe³¹; z kolei Horowitz podaje, że w szczerym zeznaniu powinno być obecnych pięć pierwszych kryteriów plus dowolne dwa dodatkowe³². Sumowanie punktów dla potrzeb analiz jest błędne i może prowadzić do błędnych decyzji³³. Innymi słowy, nie ma jasnych i obiektywnych granic, ile punktów dane zeznanie musi mieć przypisane, żeby można je było uznać za szczerze. Jedyne co można stwierdzić, wykorzystując tego rodzaju narzędzie, to fakt, że im więcej punktów, tym większe prawdopodobieństwo, że zeznanie jest prawdziwe.

Jakkolwiek te kryteria szczerości zostały stworzone na podstawie teoretycznych przesłanek, badania (zarówno terenowe, dotyczące prawdziwych spraw sądowych jak i eksperymentalne) wskazują na to, że faktycznie częściej pojawiają się w zeznaniach szczerych niż nieszczerych. Metaanaliza Oberladera i in. wskazuje na to, że w 97% badaniach terenowych oraz 65% badań eksperymentalnych kryteria CBCA częściej pojawiały się w zeznaniach szczerych niż nieszczerych³⁴. Wskazują też na to analizy wielkości efektu dla poszczególnych kryteriów. Metaanalizy wskazują na to, że wielkości efektu dla większości kryteriów w zeznaniach osób dorosłych są średnie lub wysokie³⁵, podobnie jak dla zeznań dzieci³⁶. Innymi słowy, większość kryteriów CBCA jest dobrymi lub bardzo dobrymi wskaźnikami szczerości zeznania (np. „logiczna struktura” $d = 0,48$, $d = 0,47$ u dzieci; „duża liczba szczegółów”

²⁷ R. Volbert, M. Steller, *Is this testimony truthful, fabricated, or based on false memory? Credibility assessment 25 years after Steller and Köhnken (1989)*, „European Psychologist” 2014, Vol. 19, s. 207–220.

²⁸ T.D. Boychuk, *Criteria Based Content Analysis of children's statements about sexual abuse: A field-based validation study (Unpublished doctoral dissertation)*, Arizona State University, 1991.

²⁹ Np.: H.W. Godert, M. Gamer, H.G. Rill, G. Vossel, *Statement validity assessment Inter-rater reliability of criteria-based content analysis in the mock-crime paradigm*, „Legal and Criminological Psychology” 2005, Vol. 10, s. 225–245.

³⁰ D. Bradford, *Detection of deception in the confessional context*, niepublikowana praca doktorska, 2006, University of New South Wales.

³¹ R.A. Craig, R. Sheibe, J. Kircher, D.C. Raskin, D. Dodd, *Effects of interviewer questions on children's statements of sexual abuse*, „Applied Developmental Science” 1999, Vol. 3, s. 77–85.

³² S.W. Horowitz, *Empirical support for statement validity assessment*, „Behavioral Assessment” 1991, Vol. 13, s. 293–313.

³³ Por.: K. Dużała, S.L. Sporer, R. Polczyk, *Detecting Deception: Does the Cognitive Interview Impair Discrimination with CBCA Criteria in Elderly Witnesses?* (w recenzji); V. Oberlander, C. Naefgen, J. Koppehele-Gossel, L. Quinten, R. Banse, A.F. Schmidt, *Validity of content-based techniques...*, *op. cit.*

³⁴ V. Oberlander, C. Naefgen, J. Koppehele-Gossel, L. Quinten, R. Banse, A.F. Schmidt, *Validity of content-based techniques...*, *op. cit.*

³⁵ B.G. Amado, R. Arce, F. Fariña, M. Vilariño, *Criteria-Based Content Analysis...*, *op. cit.*

³⁶ B.G. Amado, R. Arce, F. Fariña, *Undeutsch hypothesis...*, *op. cit.*

$d = 0,55$, $d = 0,77$ u dzieci). Trzeba jednak zauważyć, że są kryteria, dla których wielkość efektu jest problematyczna (np. „wybaczanie sprawcy” $d = -0,02$ dla dorosłych i $d = 0,23$ u dzieci).

Pozostaje jednak pytanie, czy CBCA jest użyteczne w kontekście sądowym? Innymi słowy, czy zwiększa trafność identyfikacji osób szczerych i nieszczerych u przeszkolonych sędziów kompetentnych? Vrij³⁷ dokonał przeglądu badań nad CBCA, wskazując, że w badaniach terenowych czułość i swoistość CBCA mogą sięgać 100%³⁸, podczas gdy eksperymentalnych zwykle balansują w okolicach 71%. Z kolei Oberlander i in.³⁹ w swojej metaanalizie wskazują na dość dużą wielkość efektu dla trafności CBCA ($g = 1,03$), która korespondowała z czułością i swoistością na poziomie 70%. Jest to znacząco więcej niż czułość i swoistość badań opartych na analizie zachowania niewerbanego⁴⁰ i porównywalnie do metod bazujących na obciążeniu poznawczym⁴¹, przy czym trzeba zaznaczyć, że badań nad podejściem poznawczym jest relatywnie niewiele w porównaniu do CBCA.

CBCA zostało zaprojektowane z myślą o zeznaniach dzieci molestowanych seksualnie – jego głównym celem była odpowiedź na pytanie: jak odróżnić zeznanie dziecka molestowanego od zeznania dziecka, które molestowane nigdy nie było? Stąd obecność w CBCA kryteriów, które mogą sprawdzać się szczególnie w przypadku dzieci (np. kryterium 10: szczegóły dobrze opisane i błędnie zrozumiałe) oraz spraw dotyczących molestowania seksualnego (np. kryterium 18: wybaczanie sprawcy). Badania eksperymentalne i terenowe pokazują jednak, że CBCA może być skutecznym narzędziem nie tylko w ocenianiu szczerości zeznań dzieci⁴², ale też dorosłych⁴³, świadków⁴⁴ oraz sprawców⁴⁵.

Jednakże trzeba pamiętać o tym, że w praktyce sama analiza CBCA nie wystarcza do trafnej oceny zeznania. Zwykle biegły powinien wziąć pod uwagę jeszcze czwarty element składający się na SVA, czyli dokonać sprawdzenia Listy Kontrolnej Prawdziwości. Procedura ta polega na ocenie sposobów i okoliczności tworzenia się zeznania. Oceny dokonuje się na 4 wymiarach: cech psychologicznych (czy język zeznania pasuje do wieku i inteligencji świadka, czy przejawia adekwatne emocje w trakcie zeznawania i jaka jest jego podatność na sugestię), cech przesłuchania (czy występowały pytania sugerujące i jakiej było ogólnie jakości), cech motywacji świadka (jakie były motywy złożenia zeznania, kontekst ujawnienia zdarzenia i czy w sprawie mogły pojawić się naciski na składanie fałszywych zeznań) oraz na wymiarze kwestii dodatkowych (czy zeznanie jest spójne z innymi zeznaniami, z prawami natury oraz z innymi dowodami). Dopiero po takiej kompleksowej ocenie zarówno treści zeznania, jak i całego kontekstu jego tworzenia można odpowiedzieć na pytanie, czy zeznanie jest szczerze, czy też nie. Niestety, w literaturze jest bardzo niewiele badań nad procedurą SVA

³⁷ A. Vrij, *Detecting lies and deceit...*, op. cit.

³⁸ Np.: P.W. Esplin, T. Boychuk, D.C. Raskin, *A field study of criteria-based content analysis of children's statement in sexual abuse cases*, paper presented at the NATO Advanced Studies Institute on Credibility Assessment, Maratea, Italy 1988, June.

³⁹ V. Oberlander, C. Naefgen, J. Koppehele-Gossel, L. Quinten, R. Banse, A.F. Schmidt, *Validity of content-based techniques...*, op. cit.

⁴⁰ C.F. Bond Jr., B.M. DePaulo, *Accuracy of deception judgments...*, op. cit.

⁴¹ A. Vrij, R. Fisher, H. Blank, *A cognitive approach to lie detection: A meta-analysis*, „Legal and Criminological Psychology” 2017, Vol. 22, s. 1–21.

⁴² P.W. Esplin, T. Boychuk, D.C. Raskin, *A field study of criteria-based...*, op. cit.

⁴³ Np.: C.L. Ruby, J.C. Brigham, *Can criteria-based content analysis distinguish between true and false statements of African-American speakers?*, „Law and Human Behavior” 1998, Vol. 22(4), s. 369–388.

⁴⁴ Np.: P. Santilla, H. Roppola, M. Runtti, P. Nemi, *Assessment of child witness statements using Criteria-Based Content Analysis (CBCA): The effects of age, verbal ability, and interviewer's emotional style*, „Psychology Crime and Law” 2000, Vol. 6(3), s. 159–179.

⁴⁵ Np.: A. Vrij, L. Akehurst, R. Soukara, R. Bull, *Detecting deceit via analyses of verbal and nonverbal behavior in adults and children*, „Human Communication Research” 2004, Vol. 30, s. 8–41.

wykonywaną w całości. Te, które są opublikowane, wskazują na bardzo dobrą stosowalność tego narzędzia w praktyce⁴⁶ i jego większą skuteczność niż CBCA używanego samodzielnie.

CBCA (a właściwie szerzej – SVA) jest obecnie dość często używane przez biegłych w praktyce sądowej. Dzięki ogromnej liczbie badań naukowych dość dobrze jest znana jego skuteczność, stosowalność oraz ewentualne przeciwwskazania do stosowania. W wielu krajach uznawane jest za jedno z najlepszych i najskuteczniejszych metod oceny zeznań i jest niemal rutynowo stosowane w sprawach dotyczących np. zeznań dzieci⁴⁷.

Innym, dość dobrze w literaturze opisanym systemem oceny treści zeznań, jest Monitorowanie Rzeczywistości (RM), oparte na bardzo dobrze zweryfikowanym i opisanym w nauce modelu Johnson i Rye⁴⁸. Zgodnie z nim wspomnienia wymyślone lub wyobrażone zawierają mniej informacji kontekstualnych (szczegółów dotyczących miejsca i czasu zdarzenia), mniej informacji sensorycznych i semantycznych i mniej emocji niż wspomnienia prawdziwe, dotyczące faktycznych zdarzeń. Dodatkowo, wspomnienia wymyślone związane są ze szczególnymi procesami poznawczymi, które nie występują przy wspomnieniach prawdziwych. Charakter i struktura tych wspomnień mogą mieć odzwierciedlenie w sposobie opowiadania o nich (czyli w zeznaniu), przez co na podstawie analizy treści zeznania można dociec, czy jest szczere, czy też nie. Narzędzie oparte na tym modelu nazywa się również Monitorowanie Rzeczywistości (RM) i ma postać kryteriów, których występowanie stwierdza się (lub nie) w zeznaniu. Kryteriów, w zależności od wersji, jest z 8 lub 7⁴⁹, lub też 43 (SMCQ⁵⁰) podzielonych na 8 grup, takich jak np. *visual details*, *spatial details*, *emotion*⁵¹. Obecnie najpowszechniej używana w badaniach jest wersja zawierająca 7 kryteriów⁵². Procedura postępowania przy RM jest taka sama, jak w przypadku CBCA – badacz czyta spisane zeznanie i ocenia, w jakim stopniu (zwykle na skali 1–6) określone kryterium pojawia się w zeznaniu. Narzędzie to używane jest głównie w badaniach naukowych i jest bardzo niewiele doniesień o jego wykorzystywaniu w praktyce.

Przegląd badań nad trafnością RM wskazuje zwykle na jego ogólną dokładność rzędu 70%⁵³. Z kolei niedawna metaanaliza⁵⁴ wskazuje na dość wysokie wskaźniki trafności ocen specjalistów, którzy dokonywali oceny zeznań przy użyciu RM. Wielkość efektu wynosiła $g = 1,03$, co przekłada się na 70% wskaźnik czułości i swoistości. Dalsze analizy wykazały, że wskaźnik ten nie był istotnie wyższy od wskaźników trafności CBCA.

Oprócz systematycznych, wystandaryzowanych narzędzi służących do wykrywania kłamstwa w oparciu o werbalne wskaźniki, w literaturze spotyka się też tzw. komputerowe analizy oparte na lingwistycznych/językowych wskazówkach kłamstwa (*linguistic cues to deception*). Są to programy komputerowe, które wyszukują w treści zeznania określone zwroty, określe-

⁴⁶ Np. w sprawach dotyczących gwałtu i fałszywych oskarżeń o gwałt: A.D. Parker, J. Brown, *Detection of deception: statement validity analysis as a means of determining truthfulness of falsity of rape allegations*, „Legal and Criminological Psychology” 2000, Vol. 5, s. 237–259.

⁴⁷ A. Vrij, *Detecting lies and deceit...*, op. cit.

⁴⁸ M.K. Johnson, C.L. Raye, *Reality monitoring*, „Psychological Review” 1981, Vol. 88, s. 67–85.

⁴⁹ Por.: S.L. Sporer, *Reality monitoring and the detection of deception*, [w:] *Deception detection in forensic contexts*, eds. P.A. Granhag, L. Stromwall, Cambridge University Press, Cambridge 2004, s. 64–102; A. Vrij, *Detecting lies and deceit...*, op. cit.; S. Sporer, *The less travelled road to truth: Verbal cues in deception detection in accounts of fabricated and self-experienced events*, „Applied Cognitive Psychology” 1997, Vol. 11, s. 373–397.

⁵⁰ S.L. Sporer, B. Kupper, *Realitätsüberwachung und die Beurteilung des Wahrheitsgehaltes von Erzählungen: Eine experimentelle Studie [Reality monitoring and the judgment of the truthfulness of accounts: An experimental study]*, „Zeitschrift für Sozialpsychologie”, 1995, Vol. 26, s. 173–193.

⁵¹ S. Sporer, *The less travelled road to truth...*, op. cit.

⁵² A. Vrij, *Detecting lies and deceit...*, op. cit.

⁵³ J. Masip, S.L. Sporer, E. Garrido, C. Herrero, *The detection of deception with the reality monitoring approach: A review of the empirical evidence*, „Psychology, Crime & Law” 2005, Vol. 11, s. 99–122.

⁵⁴ V. Oberlander, C. Naefgen, J. Koppehele-Gossel, L. Quinten, R. Banse, A.F. Schmidt, *Validity of content-based techniques...*, op. cit.

nia i frazy, które zwykle pojawiają się w zeznaniach szczerych. Obecnie najbardziej znanym programem tego typu jest LIWC (Linguistic Inquiry and Word Count)⁵⁵. Program ten analizuje spisane zeznania słowo po słowie, porównując te obecne w zeznaniach z bazą ponad 2000 słów zaklasyfikowanych do 72 kategorii, z których niektóre powiązane są z nieszczerością. Początkowo ten program zaprogramowany był po prostu do analizy treściowej zeznań, ale Newman, Pennebaker, Berry i Richards⁵⁶ po raz pierwszy wykorzystali go do wykrywania kłamstwa, uzyskując zadowalający wynik powyżej poziomu przypadku. Od tego czasu jest coraz częściej używany przez badaczy⁵⁷.

Oprócz tego istnieją również programy służące do analizy lingwistycznych wskazówek kłamstwa w nagraniach wideo (np. Agent99 Analyzer)⁵⁸ oraz wyszukujące znaczenia słów i analizujące kontekst wypowiedzi⁵⁹. Intensywna współpraca psychologów i specjalistów wykorzystujących uczenie maszynowe doprowadziła do tego, że przez ostatnie kilkanaście lat werbalne wskazówki kłamstwa stały się podstawą do stworzenia wielu różnych automatycznych algorytmów wykrywania kłamstwa⁶⁰.

Trafność takich narzędzi trudno zmierzyć, ponieważ zależy ona od bardzo wielu czynników, związanych z samą metodologią i zaprojektowaniem badania. Ostatnia metaanaliza 44 badań przeprowadzonych na 3780 osobach⁶¹ wskazała jednak, że niezależnie od programu istnieją pewne werbalne wskazówki, które wydają się silnie związane z kłamstwem. Na przykład kłamcy zawierają mniej słów w zeznaniu ($g_u = 0,33$), ich zeznania są zwykle mniej zróżnicowane pod względem słownictwa ($g_u = 0,48$), zawierają więcej słów wskazujących na gniew ($g_u = 0,27$), oraz więcej niedookreśleń i słów generalizujących ($g_u = 0,37$).

Wydaje się, że kierunek badań nad automatycznymi programami analizującymi werbalne wskazówki kłamstwa jest obiecujący. Dzięki nowoczesnym technologiom są w stanie szybko (nawet w czasie rzeczywistym) analizować wypowiedź i od razu klasyfikować zeznanie jako szczere lub nie. Do tego komputerowe algorytmy nie są podatne na często opisywany w literaturze *truth bias*, czyli charakterystyczną dla ludzi tendencję do ufania rozmówcy, co zwiększa wprawdzie swoistość, ale zmniejsza czułość wszystkich używanych narzędzi do wykrywania kłamstwa, w których to człowiek dokonuje ostatecznej decyzji. Należy jednakże pamiętać, że wszystkie tego rodzaju programy są zaprojektowane z myślą o anglojęzycznych zeznaniach. Dodatkowo walidowane są na zeznaniach określonej grupy kulturowej i są pewne przesłanki wskazujące na to, że mogą być bezużyteczne w przypadku np. subkultur czy osób posługujących się slangiem⁶².

⁵⁵ J.W. Pennebaker, M.E. Francis, R.J. Booth, *Linguistic Inquiry and Word Count (LIWC)*, Lawrence Erlbaum, Mahwah, NJ, 2001.

⁵⁶ M.L. Newman, J.W. Pennebaker, D.S. Berry, J.M. Richards, *Lying words: Predicting deception from linguistic style*, „Personality and Social Psychology Bulletin” 2006, Vol. 29, s. 665–675.

⁵⁷ Por.: V. Hauch, I. Blandón-Gitlin, J. Masip, S.L. Sporer, *Are computers effective lie detectors? A meta-analysis of linguistic cues to deception*, „Personality and Social Psychology Review” 2015, Vol. 19, s. 307–342.

⁵⁸ C.M. Fuller, D.P. Biros, J.K. Burgoon, M. Adkins, D.P. Twitchell, *An analysis of text-based deception detection tools*, In Proceedings of the 12th Americas Conference on Information Systems (3465–3472), Mexico: AIS Electronic Library, Acapulco 2006.

⁵⁹ A.C. Graesser, D.S. McNamara, M. Louwerse, Z. Cai, *Coh-Metrix: Analysis of text on cohesion and language*, „Behavioral Research Methods, Instruments, & Computers” 2004, Vol. 36, s. 193–202.

⁶⁰ J.F. Nunamaker, J.K. Burgoon, N.W. Twyman, J.G. Proudfoot, R. Schuetzler, J.S. Giboney, *Establishing a foundation for automated human credibility screening*, IEEE International Conference on Intelligence and Security Informatics, 2012, January, retrieved from: http://arizona.openrepository.com/arizona/bitstream/10150/222874/1/azu_etd_12045_sip1_m.pdf.

⁶¹ V. Hauch, I. Blandón-Gitlin, J. Masip, S.L. Sporer, *Are computers effective lie detectors?...*, *op. cit.*

⁶² *Ibidem*.

2.3. Obciążenie poznawcze (*Cognitive Load Approach*)

Podejście to opiera się częściowo na przytoczonym już modelu Zuckermana, DePaulo i Rosenthala⁶³, zgodnie z którym osoba kłamiąca doświadcza większego obciążenia poznawczego niż osoba szczerą. Vrij, Granhag i Porter zwracają uwagę, że kłamanie wymaga dużo większego wysiłku poznawczego niż zwykłe sięganie do pamięci epizodycznej w celu opisanie zdarzenia⁶⁴. Z kolei Sporer⁶⁵, odwołując się do modelu pamięci roboczej Baddeleya⁶⁶, stworzył kompleksowy model wyjaśniający różnice między kłamcami a osobami szczerymi i zwrócił uwagę, że kłamanie upośledza funkcjonowanie tzw. centralnego wykonawcy, a co za tym idzie – nie tylko wydobywanie informacji z pamięci długoterminowej, ale również kontrolowanie zachowania mimowolnego. Zgodnie z tym modelem, dodatkowe „dociążenie poznawcze” centralnego wykonawcy świadka lub podejrzanego w trakcie prowadzenia przesłuchania może znacząco uwypuklić różnice między osobami mówiącymi prawdę a kłamiącymi zarówno w sferze kontroli zachowania, jak i szybkości wypowiedzi. Zgodnie z tym podczas przesłuchania osoby kłamiące będą pokazywać oznaki przeciążenia poznawczego, po których będzie można je odróżnić od osób szczerych.

Prezentowane podejście polega zatem na zwiększaniu obciążenia poznawczego osoby przesłuchiwaną przez stosowanie różnych strategii przesłuchiwania – na przykład zadawanie zaskakujących pytań, prośby o odtworzenie wszystkich zdarzeń w odwróconej kolejności, czy wykonywanie dodatkowego zadania podczas prowadzenia przesłuchania⁶⁷. Co ważne, w ramach tego podejścia raczej nie szkoli się w żaden sposób osób wykrywających kłamstwo – nacisk jest położony na prowadzenie w odpowiedni sposób przesłuchania, aby ewentualne wskazówki przeciążenia poznawczego stały się widoczne bez specjalistycznych narzędzi i wiedzy.

Jest to stosunkowo nowe podejście i niewiele badań nad jego trafnością zostało przeprowadzonych. Dodatkowym problemem jest fakt, że strategie dociążania poznawczego są różne i rzadko się powtarzają w badaniach, a ich wpływ mierzony jest na zmiennych bardzo zależnych od rodzaju zeznania (np. liczba szczegółów związanych z napadem, niekonsekwencje w wypowiedzi itd.) Wszystko to sprawia, że bardzo trudno jest mówić o powtarzalności takiego podejścia czy sensownym zmierzeniu jego trafności. Bazując na opublikowanym przeglądzie badań⁶⁸ i metaanalizie⁶⁹ można jednak ostrożnie powiedzieć, że dwie strategie obciążania poznawczego wydają się szczególnie skuteczne. Pierwsza, zaczerpnięta z przesłuchania poznawczego⁷⁰, polega na zachęcaniu świadka lub podejrzanego różnymi sposobami, by opowiadał jak najwięcej o zdarzeniu. W obydwu publikacjach wskazano, że po zastosowaniu takiej strategii osoby nieszczerze produkowały znacząco mniej szczegółów niż osoby szczerze (nawet, jeżeli szczerze na początku deklarowały, że niewiele pamiętają ze zdarzenia). Druga strategia, która okazała się skuteczna, to zadawanie niespodziewanych pytań. Osoby,

⁶³ M. Zuckerman, B.M. DePaulo, R. Rosenthal, *Verbal and nonverbal...*, op. cit.

⁶⁴ A. Vrij, P.-A. Granhag, S. Porter, *Pitfalls and opportunities in nonverbal and verbal lie detection*, „Psychological Science in the Public Interest” 2010, Vol. 11, s. 89–121.

⁶⁵ S.L. Sporer, *Deception and cognitive load: Expanding our horizon with a working memory model*, „Frontiers in Psychology” 2016, Vol. 7, s. 420.

⁶⁶ A. Baddeley, *The episodic buffer: A new component of working memory?*, „Trends in Cognitive Sciences” 2000, Vol. 4, s. 417–423.

⁶⁷ Pełen przegląd można znaleźć w: A. Vrij, R. Fisher, H. Blank, *A cognitive approach to lie detection...*, op. cit., s. 1–21.

⁶⁸ A. Vrij, R. Fisher, H. Blank, S. Leal, S. Mann, *A cognitive approach to elicit verbal and nonverbal cues to deceit*, [w:] *Cheating, corruption, and concealment*, eds. J.W. Van Prooijen, P.A.M. Van Lange, Cambridge University Press, Cambridge 2016, p. 284–310.

⁶⁹ A. Vrij, R. Fisher, H. Blank, *A cognitive approach to lie detection...*, op. cit.

⁷⁰ R.P. Fisher, R.E. Geiselman, *Memory enhancement techniques for investigative interviewing: The cognitive interview*, Charles C. Thomas, Springfield 1992.

które kłamały, na ogół wprawdzie były przygotowane do przesłuchania, ale nie były w stanie odpowiadać tak poprawnie i szybko na niespodziewane pytania, jak osoby szczerze.

Metaanaliza Vrija i współpracowników⁷¹ wskazała na ogólną trafność podejścia opartego na zwiększaniu obciążenia poznawczego wynosząca 71% (w porównaniu do 56%, kiedy żadne narzędzie nie zostało zastosowane). Co ciekawe, osoby wykrywające kłamstwo w tych badaniach nie były w żaden sposób szkolone czy pouczane – bazowały tylko na swojej intuicji. Różnica brała się tylko z tego, że osoby kłamiące były przesłuchiwane albo przy użyciu tradycyjnego przesłuchania policyjnego, albo przy użyciu strategii zwiększających obciążenie poznawcze. Prawdopodobnie dodatkowe szkolenie dotyczące wyszukiwania sygnałów takiego obciążenia mogłoby jeszcze bardziej zwiększyć trafność.

Podejście to wydaje się obiecujące, należy jednak pamiętać że ma swoje ograniczenia. Na przykład jest wiele okoliczności, w których kłamstwo nie będzie wywoływać obciążenia poznawczego⁷². Ponadto podejście to nie ma zastosowania w niektórych rodzajach kłamstwa, takich jak kłamstwa dotyczące intencji⁷³. Wydaje się więc, że potrzebnych jest jeszcze wiele badań, żeby można było stosować to podejście w praktyce.

3. Przyszłość

Rozwój cywilizacji stawia nowe wyzwania przed badaczami zajmującymi się nieinstrumentalnymi metodami detekcji kłamstwa. Zagrożenie terroryzmem implikuje konieczność znalezienia metod, które pozwolą na wykrywanie kłamstw dotyczących nie tylko dokonanych czynów lub wspomnień, ale intencji⁷⁴. Z drugiej strony rozwój nowoczesnych technologii komunikacyjnych (portale społecznościowe, smsy, posty) inspirowała badaczy zajmujących się lingwistycznymi i treściowymi metodami do znalezienia jeszcze lepszych narzędzi do wykrywania kłamstwa w krótkich wiadomościach tekstowych⁷⁵. Bardzo niewiele obecnie istniejących metod może być stosowana, kiedy świadkami są osoby z zaburzeniami psychicznymi lub o obniżonym funkcjonowaniu poznawczym – obecne prace nad ulepszeniem metod nieinstrumentalnych zmierzają do tego, by przystosować je do takich przypadków⁷⁶.

Jaka jest zatem przyszłość nieinstrumentalnych metod? Wydaje się, że nawet przy dość prężnie rozwijających się metodach instrumentalnych (takich jak wykrywanie kłamstwa w oparciu o EEG, termowizję czy fMRI) metody nieinstrumentalne będą wciąż rozwijane i wykorzystywane. Ich niewątpliwą zaletą jest relatywna łatwość aplikowania i brak konieczności posiadania specjalistycznej, nieraz bardzo drogiej aparatury. Ich wykorzystanie jest też szybsze i łatwiejsze. Z drugiej jednak strony ta łatwość aplikowana może być pozorna – na przykład metod treściowych nie powinno się używać, jeżeli protokół z przesłuchania nie jest dokładną transkrypcją zeznania, RM ma zastosowanie tylko do całości zeznania. Dodatkowo w praktyce wielu biegłych, wiedzionych pozorną łatwością wykorzystania np. CBCA, może korzystać z niego bez przeszkolenia, co prowadzi do znacznego spadku trafności i rzetelności tego narzędzia⁷⁷. Należy też pamiętać, że nie wszystkie metody nieinstrumentalne mają po-

⁷¹ A. Vrij, R. Fisher, H. Blank, *A cognitive approach to lie detection...*, op. cit.

⁷² S.L. Sporer, *Deception and cognitive load...*, op. cit.

⁷³ E. Fenn, M. McGuire, S. Langben, I. Blandón-Gitlin, *A reverse order interview does not aid deception detection regarding intentions*, „Frontiers in Psychology” 2015, Vol. 6, s. 1298.

⁷⁴ P.A. Granhag, E. Mac, *Preventing future crimes: identifying markers of true and false intent*, „European Psychology” 2014, Vol. 19, s. 195–206.

⁷⁵ Np.: M.E. Smith, J.T. Hancock, L. Reynolds, J. Birnholtz, *Everyday deception or a few prolific liars? The prevalence of lies in text messaging*, „Computers in Human Behavior”, 2014, Vol. 41, s. 220–227.

⁷⁶ Np.: A.L. Manzanero, A. Alemany, M. Recio, R. Vallet, J. Aróztegui, *Evaluating the credibility of statements given by persons with intellectual disability*, „Anales de Psicología” 2015, Vol. 31, s. 338–344.

⁷⁷ V. Hauch, I. Blandón-Gitlin, J. Masip, S.L. Sporer, *Are computers effective lie detectors?...*, op. cit.

twierdzoną skuteczność. O ile SVA czy RM są metodami bardzo dobrze przebadanymi, o tyle np. podejście polegające na przeciążeniu poznawczym jest dopiero w fazie badań, a z kolei metody oparte na analizie niewerbalnego zachowania są nieskuteczne, ale powszechnie stosowane – zwłaszcza przez policjantów czy prokuratorów⁷⁸. Wydaje się zatem, że rolą badaczy jest nie tylko poszukiwanie nowych, bardziej skutecznych nieinstrumentalnych narzędzi służących do wykrywania kłamstwa, ale również popularyzacja wiedzy o już istniejących.

4. Kwestie prawne związane z możliwością wykorzystania nieinstrumentalnych metod detekcji kłamstwa

Co najmniej kilka kwestii prawnych wymaga tu rozważenia. Kwestia pierwsza to wykorzystanie nieinstrumentalnych technik w polskim procesie karnym. Tu mamy do rozwiązania dwa podstawowe problemy. Po pierwsze, kto tymi technikami miałby się posługiwać: organ procesowy czy biegły?

Jeśli organ procesowy (czyli w praktyce prawnik, bez większego psychologicznego przygotowania), to jaki byłby walor końcowy takiej czynności? Dowód (na co?), nabranie określonego przekonania przez organ procesowy? Jakie miałyby być konsekwencje procesowe i faktyczne takiego przekonania?

Jeśli biegły, to czy w formie odrębnej ekspertyzy (na podstawie art. 193 kpk), czy poprzez udział w czynności prowadzonej przez organ procesowy (udział w przesłuchaniu na podstawie art. 192 §2 kpk, art. 185a §2 kpk). Jeśli w formie odrębnej ekspertyzy, to jak należałoby określić zadanie biegłego?

Czy przy relatywnie niskiej wartości diagnostycznej takich metod, wynik ekspertyzy miałby walor dowodu?

Sensowna analiza treści zeznania, dokonywana na podstawie analizy treści protokołu tej czynności, wymagałaby co najmniej ujednolicenia sposobu sporządzania tego protokołu. Kodeks postępowania karnego w art. 143 §1 pkt 2 nakazuje spisanie protokołu z „przesłuchania oskarżonego, świadka, biegłego i kuratora”, w art. 143 §1 pkt 10 – z „przebiegu posiedzenia sądu, jeżeli stawia się na nim uprawnione osoby albo ich obecność jest obowiązkowa” a także w art. 143 §1 pkt 11 – z „przebiegu rozprawy”.

Przebieg czynności protokołowanych „może być utrwalony ponadto za pomocą urządzenia rejestrującego obraz lub dźwięk” (art. 147 kpk). W sytuacji gdy czynność jest rejestrowana w ten sposób, protokół pisemny można ograniczyć „do zapisu najbardziej istotnych oświadczeń osób biorących w niej udział” (art. 147 §3 kpk). Z kolei art. 171 kpk stanowi, że „osobie przesłuchiwanej należy umożliwić swobodne wypowiedzenie się w granicach określonych celem danej czynności, a dopiero następnie można zadawać pytania zmierzające do uzupełnienia, wyjaśnienia lub kontroli wypowiedzi” (art. 171 §1 kpk). Przy czym pytania mogą, zadawać oprócz organu procesowego, strony, obrońcy, pełnomocnicy, biegli (art. 171 §2 kpk).

Jak już z przytoczonych przepisów wynika, na treść, objętość, szczegółowość zeznań (wyjaśnień) i czas ich trwania ma wpływ wiele czynników. Aktywność osoby przesłuchiwanej (determinowana także przez wiele czynników, w tym niewątpliwie jej osobowość, inteligencję, zdolność spostrzegania etc.) jest tylko jednym z nich. Wiele zależy od aktywności organu procesowego oraz osób uprawnionych do zadawania pytań, ich dociekliwości, poziomu aspiracji, znajomości materii etc.

⁷⁸ L.H. Colwell, H.A. Miller, R.S. Miller, P.M. Lyons, *US police officers'...*, *op. cit.*

Co więcej, przepisy kpk nie precyzują sposobu, w jaki ma być tworzony protokół. Praktyka jest tu bardzo różna. W postępowaniu przygotowawczym, protokół jest pisany albo osobiście przez przesłuchującego, albo przez protokolanta. Przed sądem – zawsze przez protokolanta. Zwykle słowa przesłuchiwanego do protokołu dyktuje organ procesowy (prokurator, policjant, sędzia). Rzadziej protokolant pisze sam „ze słuchu”. Dyktujący protokół czasem powtarza słowo w słowo treść zeznania (wyjaśnienia), najczęściej jednak upraszcza wypowiedź, koncentrując się na sprawach najważniejszych, poprawiając przy tym błędy stylistyczne, gramatyczne czy językowe przesłuchiwanego. Czasem w protokole odnotowuje się dosłownie zadane pytania i udzielone na nie odpowiedzi, czasem (w sądzie jest to reguła) nie odnotowuje się treści pytań, a najwyżej fakt, że jest to odpowiedź na pytanie sądu, oskarżyciela czy obrońcy (bez podawania treści tego pytania). Często fakt zadawania pytań w ogóle nie jest odnotowywany, a zapisuje się jedynie wypowiedzi osoby przesłuchiwanej. W takiej sytuacji, nie wiadomo, co było wypowiedzią spontaniczną, co – odpowiedzią na pytanie. Wtedy analiza treści protokołu pod kątem oceny prawdomówności przesłuchiwanego (np. metodą RM czy CBCA/SVA) nie jest możliwa.

Można postulować bardziej szczegółowe unormowanie sposobu sporządzania protokołu z przesłuchania, co najmniej przez wprowadzenie obowiązku wpisywania treści zadanych przesłuchiwanemu pytań, a także obowiązku nagrywania przebiegu tej czynności. Umożliwiłoby to powszechniejsze wykorzystanie metod analizy treści do oceny wiarygodności zeznania (wyjaśnienia).

Część V

Próba wykorzystania zmian temperatury twarzy jako wskaźnika zmian emocjonalnych przy detekcji kłamstwa

W świetle aktualnego stanu wiedzy nie ulega wątpliwości, że zmiany temperatury twarzy są dobrym wskaźnikiem zmian emocjonalnych. Kamera termowizyjna niewątpliwie znacznie ułatwia pomiar tej temperatury i jej zmiany. Mierzenie temperatury w poszczególnych partiach twarzy, śledzenie i rejestracja tych zmian metodami tradycyjnymi były znacznie utrudnione. Przegląd literatury na temat rejestracji zmian temperatury twarzy za pomocą kamery termowizyjnej prezentuje artykuł M. Gołaszewskiego, P. Zająca i J. Widackiego: *Thermal Vision as a Method of Detection of Deception. A Review of Experiences*, „European Polygraph” 2015, 9, 1 (31), s. 5–24 (por. niżej).

Oczywiście zmiany temperatury twarzy są tylko jednym z wskaźników zmian emocjonalnych. Warto przypomnieć, że tradycyjny poligraf rejestrował co najmniej trzy takie wskaźniki (zmiany w przebiegu czynności oddychania, zmiany w aktywności skórno-galwanicznej (reakcji skórno-galwanicznej) oraz w przebiegu pracy układu krążenia (zmiany w częstotliwości tętna, względne wahania ciśnienia krwi). Jednak gdyby dobrać trzy inne wskaźniki zmian emocjonalnych (np. zmiany temperatury twarzy, zmiany głosu i ruchy gałek ocznych), można by teoretycznie uzyskać efekt identyczny jak przy badaniu poligraficznym, a nie wymagałoby to zakładania na ciele badanego żadnych czujników, czyli teoretycznie można by takie badanie wykonać bez wiedzy, a zatem także bez zgody badanego. Rodziłoby to nowe problemy natury prawnej i etycznej. Do tych problemów odniesiemy się w części VI. Jednak rejestracja zmian temperatury twarzy pod wpływem bodźców emocjonalnych, dokonywane nawet za pomocą kamery termowizyjnej, rodzi szereg trudności natury technicznej. Próbowaliśmy te problemy rozwiązać w badaniach eksperymentalnych, opisanych w kolejnych artykułach. Pierwszym problemem było przekształcenie obrazu z kamery termowizyjnej, w postaci zmieniających barwy plam, w wykres. Częściowo ten problem udało się nam rozwiązać, co opisujemy w artykule J. Widacki, M. Widacki, J. Antos, *Preparation to Experimental Testing of the Potential from Using Facial Temperature Changes Registered with an Infrared Camera in Lie Detection* („European Polygraph” 2016, 10, 1 (35) 17–27, por. niżej). Artykuł ten wzbudził spore zainteresowanie, czego dowodem był m.in. jego przedruk w języku rosyjskim¹. Okazało się, że nie jest to jedyny problem. Kolejnym było to, że na skutek ruchów

* Dr inż. Zbigniew Mikrut, AGH Akademia Górniczo-Hutnicza, Wydział Elektrotechniki, Automatyki, Informatyki i Inżynierii Biomedycznej.

¹ J. Widacki, M. Widacki, J. Antos, *Подготовка к проведению полиграфного исследования с использованием тепловизора путем измерения динамики изменения температуры лица*, [w:] *Актуальні питання теорії та практики використання поліграфа*, Київ 2016, s. 81–88.

głowy badanego, a nawet ruchów mięśni twarzy spowodowanych wypowiedaniem przez badanego krótkich odpowiedzi na pytania badającego („tak” lub „nie”), obserwowany przez kamerę fragment twarzy znikał z pola jej widzenia. Wymagało to opracowania nowego programu komputerowego i nowych eksperymentów (por. Z. Mikrut, M. Widacki, J. Widacki, *Próba wykorzystania zmiany... w niniejszym tomie*).

W świetle naszych doświadczeń, a także w świetle doświadczeń innych autorów, metoda ta, choć bardzo obiecująca, nie nadaje się jeszcze do rutynowego stosowania (por. część VI). Jeśli zaś idzie o wykorzystanie tej metody do monitorowania tłumu (np. na lotniskach, w czasie imprezach masowych), to skuteczniejsza wydaje się, przy obecnym stanie techniki, metoda obserwacji mimiki i innych ruchów wyrazowych poszczególnych osób z tłumu na obrazie ze zwykłej kamery².

² О. Алесковская, *Профайлинг и его перспективное развитие*, [w:] *Применение полиграфа в бывшем восточном блоке*, под редакцией Я. Видацкого, В. Шаповалова, И. Усикова, Офісна Wydawnicza AFM, Kraków 2017, s. 195–201.

Zbigniew Mikrut, Michał Widacki, Jan Widacki

Próba wykorzystania zmiany temperatury twarzy jako wskaźnika w instrumentalnej detekcji kłamstwa

1. Wprowadzenie

Zmiany temperatury twarzy są niewątpliwie dobrym wskaźnikiem zmian emocjonalnych i dlatego podejmowane są próby wykorzystania ich w detekcji kłamstwa¹. Zmiany te mogą być obserwowane i rejestrowane przez kamerę termowizyjną, a więc na odległość, bez konieczności zakładania na ciało badanego jakiegokolwiek czujnika, co powoduje, że metoda ta wydaje się atrakcyjna dla detekcji kłamstwa. Teoretycznie pozwala ona na zastosowanie jej bez wiedzy, a więc tym bardziej zgody badanego, co może mieć znaczenie w działaniach służb policyjnych lub specjalnych.

Zmiany temperatury twarzy w obrazie termowizyjnym manifestują się zmianami barw różnych jej części. Mogą to być barwy w pełnej skali od niebieskiej do ciemnoczerwonej.

Jest co najmniej kilka możliwych metod przekształceń zmieniających się plam barwnych w dający się rejestrować wykres graficzny. Jedna z nich została opatentowana w USA i Anglii², inna – w Anglii³. Pierwszą naszą samodzielną udaną próbę takiego przetworzenia zmieniających barwy plam w wykres graficzny opisaliśmy w 2016 roku⁴.

Po rozwiązaniu tego problemu pojawił się problem kolejny. Niewielkie nawet ruchy głowy osoby badanej, tak małe, że byłyby nieistotne i nie utrudniałyby klasycznego badania poligraficznego, powodowały, że obserwacja wybranych do monitorowania fragmentów twarzy byłaby przerywana, bowiem obserwowany fragment znikał z pola widzenia kamery. Zaistniała zatem potrzeba opracowania programu, który pozwoliłby kamerze śledzić wybrany fragment twarzy w sposób ciągły.

Realizacja odpowiedniego algorytmu oraz wyniki testu zostały opisane w ustępie 2.

¹ Przegląd tej literatury dokonany został w pracy: M. Gołaszewski, P. Zając, J. Widacki, *Thermal Vision as a Method of Detection of Deception*, „European Polygraph” 2015, 9, 1 (31).

² Patent US 6996256, patent EP 1286620 (Detection system and method using thermal image analysis, 2006).

³ Patent EP 0885587 (Thermal imaging method and apparatus, 1998).

⁴ J. Widacki, M. Widacki, J. Antos, *Preparation to Experimental Testing of the Potential from Using Face Temperature Changes Registered with an Infrared Camera in Lie Detection*, „European Polygraph” 2016, 10, 1 (35).

W ustępie 3 opisano wyniki obserwacji zmian temperatury dla kilku prostokątnych obszarów („okien”) zlokalizowanych w różnych obszarach twarzy i następnie porównano je z rejestracją poligraficzną obejmującą przebieg czynności oddychania, przebieg pracy układu krążenia (częstotliwość tętna i względne wahania ciśnienia krwi), oraz reakcję skórno-galwaniczną (elektrodermalną). Zbadano też wpływ rozmiaru i kształtu wybranego „okna” na otrzymane wyniki.

Rejestracje reakcji osoby badanej poligraficznie przy użyciu techniki Utah Zone Comparison Test (Utah ZCT) posłużyły do sprawdzenia algorytmu wykrywania zmian temperatury twarzy w warunkach stosunkowo znacznego (ok. 200 cm) oddalenia kamery od twarzy osoby badanej (por. ust 3).

2. Eliminacja wpływu niewielkich ruchów twarzy osoby badanej

Realizacja tego celu wymagała w pierwszej kolejności wyznaczenia „okien”, czyli prostokątnych pól na twarzy osoby badanej, w których ma być prowadzona analiza obrazu⁵. Konieczne było dokonanie wyboru obszaru („okna”), który w sposób ciągły śledzony będzie przez kamerę termowizyjną na kolejnych klatkach filmu. Wybór takiego obszaru musiał uwzględniać równocześnie trzy kryteria. Po pierwsze, obszar musi zawierać obiekty kontrastowe pod względem barwy. Po drugie, być położony w rejonie, gdzie zmiany temperatury są największe, czyli w sąsiedztwie dużych naczyń krwionośnych. Po trzecie wreszcie, objęte nim obiekty powinny być stałe co do kształtu i nie zmieniać swego położenia względem „okien”, a tym samym względem kamery, czyli powinny to być obszary najmniej zmieniające się pod względem ruchów mimicznych. Uwzględniając te kryteria wybrano jako „okno” wzorcowe dolną część nosa. Wymagało to takiego ustawienia kamery, by widoczne były dziurki nosa. Okno wzorcowe zostało otoczone ramką, w której w kolejnych klatkach filmu poszukiwany był wzorec. Rozmiar ramki wyznaczył zakres dopuszczalnych ruchów twarzy. Wzorec był przesuwany wewnątrz tak zakreślonej ramki. W każdym nowym położeniu obliczany był współczynnik korelacji:

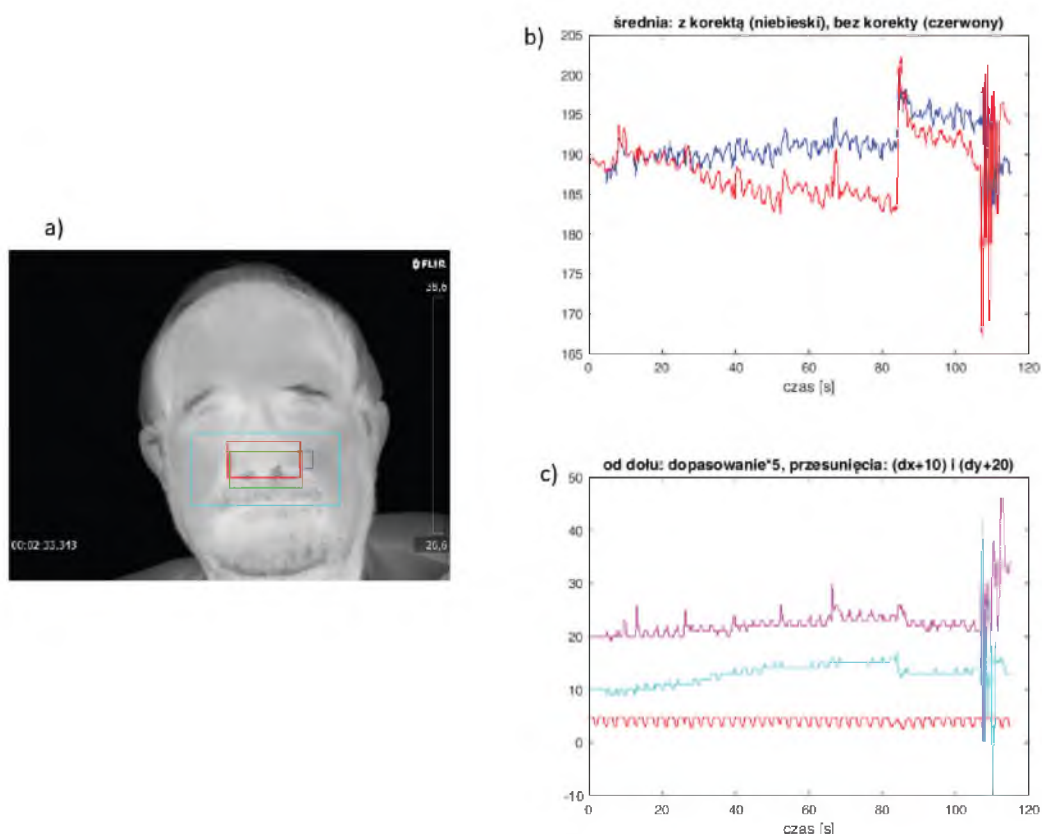
$$\gamma(u, v) = \frac{\sum_{x,y} [f(x, y) - \bar{f}_{u,v}] [t(x - u, y - v) - \bar{t}]}{\left\{ \sum_{x,y} [f(x, y) - \bar{f}_{u,v}]^2 \sum_{x,y} [t(x - u, y - v) - \bar{t}]^2 \right\}^{0.5}}$$

gdzie f jest obrazem ramki, w której poszukiwany jest wzorec t , $\bar{f}_{u,v}$ jest wartością średnią obrazu $f(x,y)$ w obszarze pod wzorcem t , \bar{t} jest wartością średnią obrazu wzorca.

Proces ten był realizowany przy pomocy funkcji w programie Matlab o nazwie „normcorr2”⁶. Obliczone współczynniki $\gamma(u, v)$ były zapisywane w macierzy. Współrzędne maksymalnego elementu macierzy odpowiadały najlepszemu dopasowaniu. Proces dopasowywania przedstawiono na rycinie 1.

⁵ MATLAB release 2016b, The Math Works, Inc. Natick, Massachusetts.

⁶ J.P. Lewis, *Fast Template Matching*, Vision Interface, 1995, s. 120–123.



Ryc. 1. Proces dopasowywania obszarów i korekty położenia śledzonego obiektu (opis w tekście)

Na obrazie przedstawionym na rycinie 1a zaznaczono obszary wykorzystywane w algorytmie.

- Ramka zielona oznacza początkowe położenie nosa – będącego obszarem wzorca, który dopasowywany był w kolejnych kadrach filmu.
- Ramka jasnoniebieska określa obszar poszukiwań wzorca.
- Ramka czerwona otacza najlepsze uzyskane dopasowanie.
- Ramka ciemnoniebieska wyznacza obszar, z którego obliczana była średnia temperatura, będąca odzwierciedleniem wartości średniej, obliczonej dla pikseli zawierających się w ramce.

Przebieg zmian temperatury w opisanych powyżej oknach („ramkach”) w badanym obszarze ilustruje rycina 1b. Kolorem czerwonym oznaczono wartości średnie obliczone dla pikseli z ramki ciemnoniebieskiej, która nie zmieniła swego położenia. Wykres w kolorze niebieskim to analogiczny przebieg uwzględniający korektę położenia, związaną z ruchami twarzy, z ramki ciemnoniebieskiej, której położenie ściśle związane było ze znalezionym punktem dopasowania wzorca (ramka czerwona). Korekta położenia spowodowała na wykresie niebieskim (ryc. 1b) widoczną zmianę średniej temperatury w tym samym rejonie twarzy, o czym świadczy poziomy charakter linii trendu wykresu niebieskiego.

Na rycinie 1c przedstawiono dodatkowe wykresy ilustrujące przebieg procesu dopasowywania. Wykresy zostały przesunięte w pionie lub tak przeskalowane, aby nie nakładały się na siebie. Dolny wykres przedstawia kolejne wartości współczynników dopasowania \mathcal{V} (u , v). Zostały one przemnożone przez 5. Widoczne na tym wykresie oscylacje odzwier-

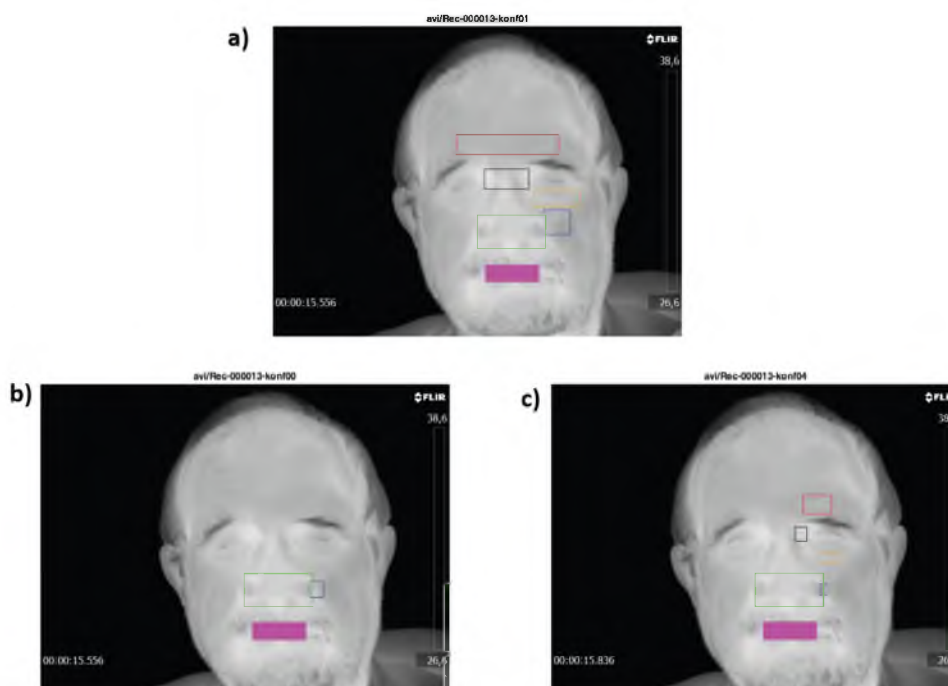
ciadają niewielkie zmiany na kolejnych obrazach będące wynikiem ruchów twarzy spowodowanych oddychaniem badanego przez nos. Dwa kolejne wykresy prezentują zarejestrowane zmiany położenia twarzy („punktu dopasowania”) odpowiednio na osi x (kolor niebieski – cyan) i osi y (kolor wiśniowy – magenta). Jak widać, są to odchylenia o wartości kilku pikseli (oś pionowa dla dwóch ostatnich wykresów odpowiada odległościom w pikselach). Większe zmiany, widoczne przy końcu badania, są efektem celowych szybkich ruchów głowy badanego.

3. Rejestracja poligraficzna a wyniki detekcji termowizyjnej

Omawiany eksperyment prowadzono z użyciem kamery termowizyjnej FLIR A655sc, która była ustawiona na statywie w odległości ok. 90 cm od twarzy osoby badanej. Ustawienie kamery było takie, by obraz twarzy osoby badanej wypełniał większą część kadru. Za pomocą dostarczonego przez producenta oprogramowania rejestrowano filmy o rozdzielczości 640x480 pikseli, z częstotliwością 30 klatek na sekundę, które następnie przetworzono do postaci akceptowanej przez oprogramowanie Matlab (w wersji 2016b).

Nagrania filmu dokonano podczas wykonywania realizacji testu stymulacyjnego poprzedzającego rutynowe badanie poligraficzne⁷, w czasie którego zadano osobie badanej 6 pytań (bodźców).

Film rejestrujący zmiany obrazu twarzy osoby badanej analizowano za pomocą trzech konfiguracji, przedstawionych na rycinie 2. Podobnie jak poprzednio ramka zielona ogranicza obszar nosa, wykorzystywany do obliczania metodą korelacji niewielkich przesunięć spowodowanych ruchami głowy badanego. Pozostałe obszary („ramka”) wyznaczono na podstawie analizy przebiegu naczyń krwionośnych twarzy⁸. Ramka czerwona służyła do określenia zmiany położenia ust.



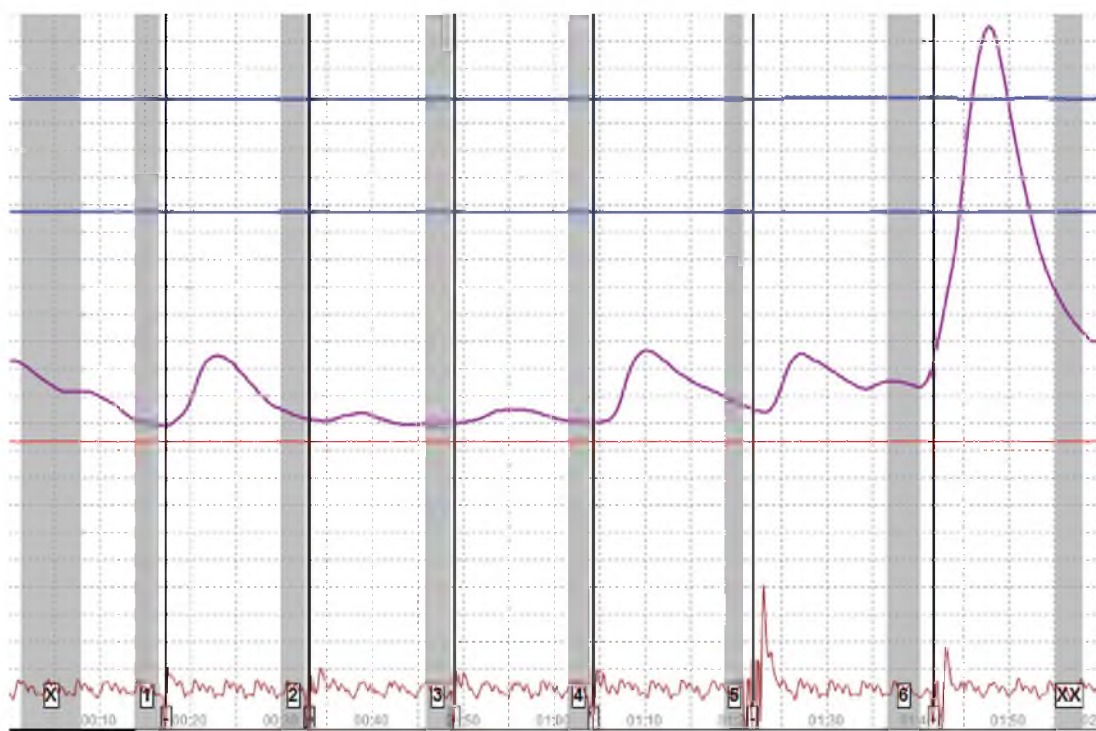
Ryc. 2. Trzy przykłady konfiguracji: a); b); c).

⁷ Por. *Kryminalistyka*, red. J. Widacki, wyd. 3, CH Beck, Warszawa 2016.

⁸ Por. R. Aleksandrowicz, *Mały atlas anatomiczny*, PZWL, Warszawa 2009, s. 97.

Konfiguracja „konf00” składała się z jednego obszaru średniej wielkości zlokalizowanego w okolicy nosa (por. ryc. 2b). Dwie pozostałe (konf01 i konf04) umożliwiają pobieranie informacji z obszarów czoła, nasady nosa, policzka pod lewym okiem, policzka koło nosa. Obszary te wybrano, kierując się wspomnianymi wyżej trzema kryteriami. Różne są rozmiary wybranych obszarów: dużo większe w przypadku „konf01” niż w „konf04” (por. ryc. 2a i 2c).

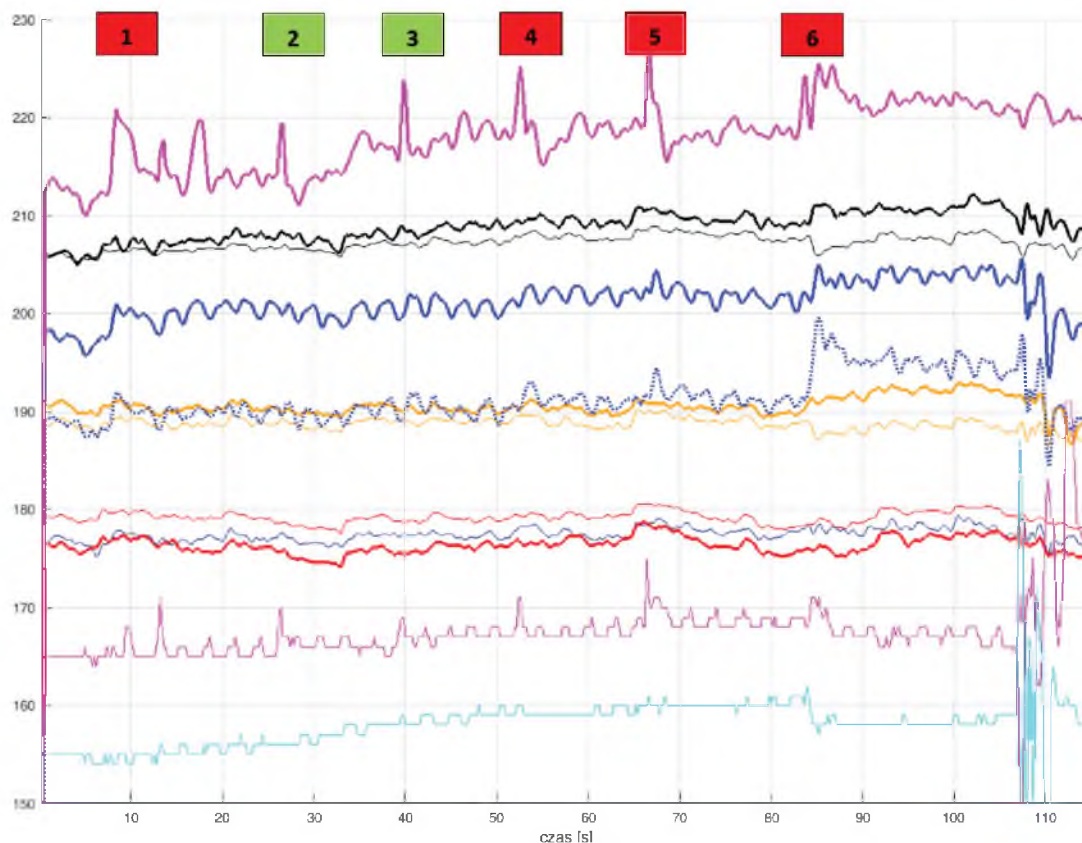
Na ryc. 3 przedstawiono fragment zapisu poligraficznego z wynikiem rejestracji czujnikiem psychogalwanometru, rejestrującego zmiany oporności skóry osoby badanej (odruch skórno-galwaniczny, *galvanic skin response*, GSR, reakcja elektrodermalna, *electrodermal activity*, EDA), będącego częścią składową poligrafu Lafayette LX-4000.



Ryc. 3. Wynik rejestracji czujnikiem GSR w czasie testu stymulacyjnego

Rycina 4 przedstawia wykresy dla omówionych wyżej konfiguracji. Momenty prezentacji bodźca zaznaczono kolorowymi prostokątami u góry ryciny. Kolor zielony prostokąta oznacza sygnał odebrany na krzywej GSR, którego odpowiednika będziemy poszukiwać na wykresie termo. Kolory wykresów przedstawionych na poniższej rycinie odpowiadają kolorom obszarów ukazanych na ryc. 2. Konfigurację „konf.00” przedstawiono linią kropkowaną, „konf.01” – liniami cienkimi, „konf04” – liniami pogrubionymi.

Analiza wizualna, przeprowadzona na podstawie oceny wysokości maksimów lokalnych, pokazuje istotne różnice, obserwowane dla różnych położeń obszarów i różnych wielkości. Zestawienie to, dla poszczególnych odpowiedzi, przedstawia tabela 1.



Ryc. 4. Średnie obliczone w oknach zdefiniowanych w konfiguracjach nr: 00 – linia kropkowana, 01 – linie cienkie, 04 – linie grube. Od dołu: ruchy twarzy w kierunku x (jasnoniebieski), w kierunku y (wiśniowy). U góry: obszar ust (linia wiśniowa, gruba). Do funkcji obliczonych w oknach zastosowano filtr uśredniający o rozmiarze 3. Ponumerowane prostokąty to odcinki czasu, w których badana osoba udzielała odpowiedzi.

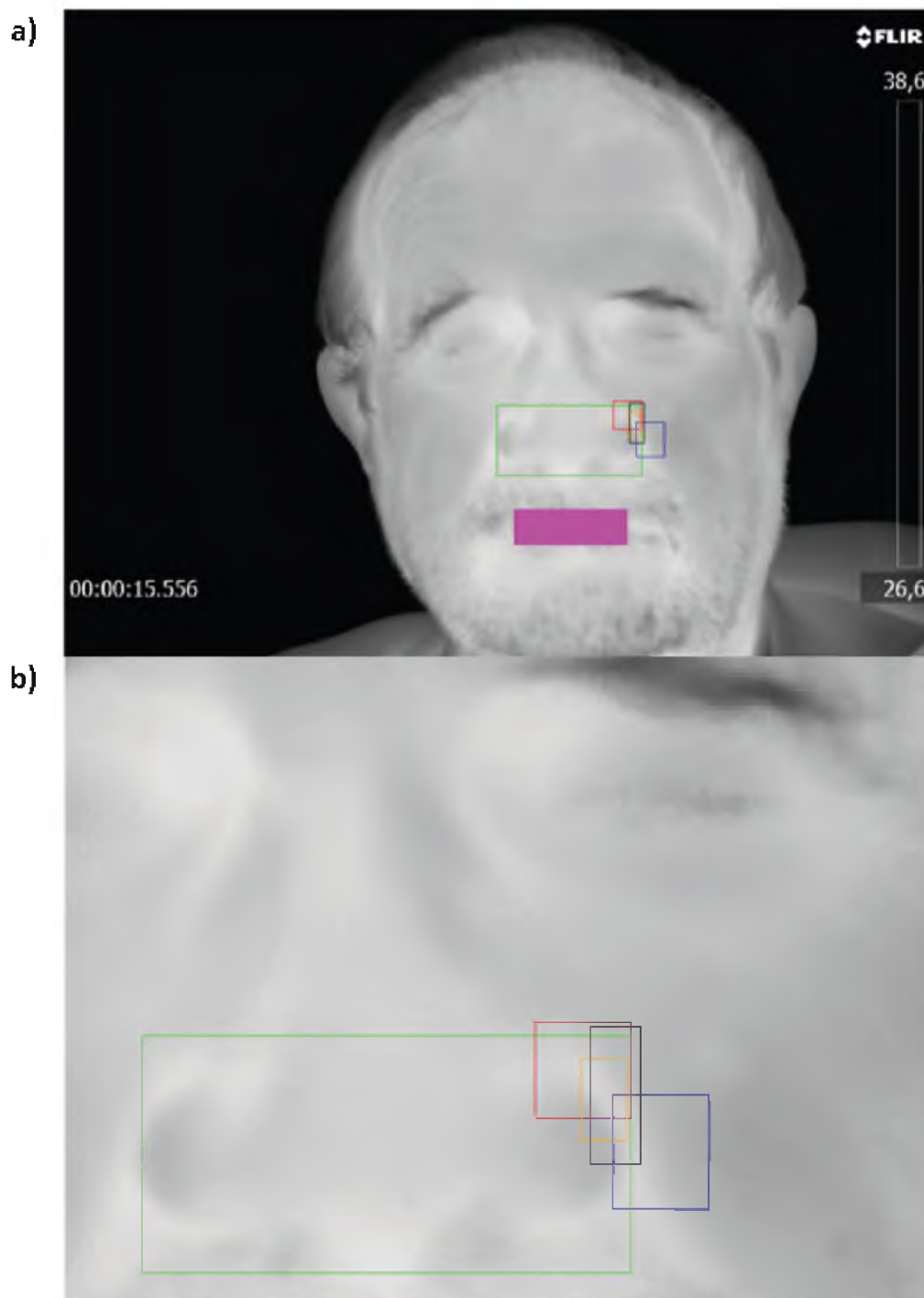
Tabela 1. Analiza wizualna dla zdefiniowanych obszarów: wystąpienie maksimów lokalnych

Konfig.	Obszar	Rozmiar	odpowieź						uwagi/(ogólna ocena)
			1	2	3	4	5	6	
konf01	czoło	Duży	+	+	-	-	+	+	niektóre odpowiedzi są opóźnione/(-)
konf04	czoło	Mały	+	+	-	-	+	+	
konf01	nasada nosa	Duży	+	-	-	-	+	-!	minimum zamiast max/(-!)
konf04	nasada nosa	Mały	+	-	-	-	+	+	(-)
konf01	policzek pod okiem	Duży	+	-	-	-	+	-!	minimum zamiast max/(-!)
konf04	policzek pod okiem	Mały	+	-	-	-	+	+	(-)
konf01	policzek koło nosa	Duży	+	+	-	+	+	-	(-)
konf04	policzek koło nosa	Mały	+	-	-	+	+	+	(+)
konf00	policzek koło nosa	Średni	+	-	-	+	+	+	(+)

Jak to pokazują ryc. 4 i tabela 1, linie dla obszarów dużych i małych są zgodne, jednak nie umożliwiają różnicowania między reakcjami towarzyszącymi odpowiedziom prawdziwym i fałszywym. Sytuacje nietypowe zaznaczono wykrzyknikami. Zwraca uwagę, że amplitudy otrzymanych funkcji są niższe dla obszarów większych. Pełną zgodność otrzymano dla obszarów opisanych w dwóch ostatnich wierszach tabeli 1.

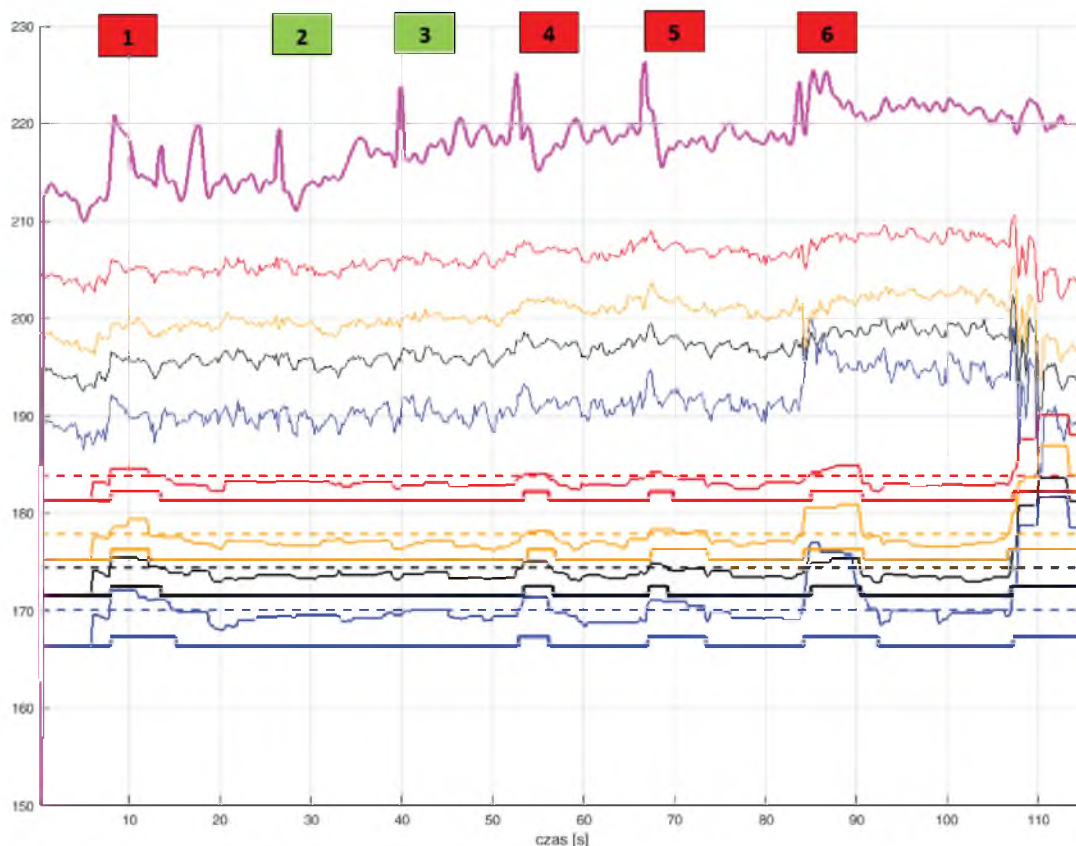
4. Wpływ niewielkich zmian położenia i rozmiarów okien analizy dla jakości zapisu

Na ryc. 5 przedstawiono konfigurację „konf05” wykorzystaną w tej analizie. Obszarem bazowym była „ramka” określona w konfiguracji „konf00” (por. ostatni wiersz tabeli 1 oraz ryc. 2b). Kolejne obszary analizy przesuwano względem obszaru bazowego w kierunku czoła osoby badanej, zmieniając ich rozmiar (por. ryc. 5).



Ryc. 5. a) Widok ogólny konfiguracji „konf05”, b) powiększenie fragmentu ze zdefiniowanymi obszarami

Na ryc. 6 zestawiono funkcje (wykresy) uzyskane, tak jak poprzednio, metodą uśredniania wartości pikseli poszczególnych obszarów. Pięć funkcji (licząc od góry) zaprezentowano w postaci „surowej”, co oznacza, że nie zostały one poddane filtracji uśredniającej. Kolejność prezentacji funkcji względem osi pionowej odpowiada kolejnym położeniom dolnych krawędzi obszarów zdefiniowanych w konfiguracji „konf05”, a przedstawionych na ryc. 6.



Ryc. 6. Od góry: funkcja odpowiadająca obszarowi 'usta', poniżej: cztery funkcje obliczone w obszarach konfiguracji „konf05” (bez filtracji uśredniającej), dalej: wyniki detekcji i binaryzacji lokalnych maksimów czterech funkcji (opis w tekście)

Pomimo różnic w położeniu obszarów i ich rozmiarów, funkcje mają zbliżony kształt. Jedyne znaczące różnice pojawiły się w okolicy 85 sekundy rejestracji.

Detekcja maksimów odpowiadających niezgodnym z prawdą odpowiedziom badanego jest w przypadku tak przedstawionych wykresów zadaniem trudnym. Dlatego też zdecydowano się na zastosowanie wspomagającego tę analizę algorytmu. W przesuwającym się wzdłuż danej funkcji oknie, obliczana była amplituda. Amplitudę obliczano w ten sposób, że od wartości maksymalnej odejmowano wartość minimalną. Powstałe w ten sposób funkcje (wykresy) zostały przedstawione w dolnej części ryc. 6 (w zakresie od 165 do 185 wartości osi pionowej).

Tą metodą zrealizowano zadanie uwypuklenia maksimów lokalnych. Dzięki temu możliwe było przeprowadzenie linii (progów – na ryc. 6 są to linie przerywane) lokalizujących te maksima, a tym samym wykrywających efekt zmiany temperatury spowodowane zmianami emocjonalnymi występującymi u osoby badanej.

Położenie linii lokalizujących maksima (progów) można wyznaczyć na wykresach ręcznie. Podjęto jednak próbę automatycznej generacji ich położenia, jak się okazało, w większości przypadków zakończonej powodzeniem. Wypróbowano w tym celu trzy metody:

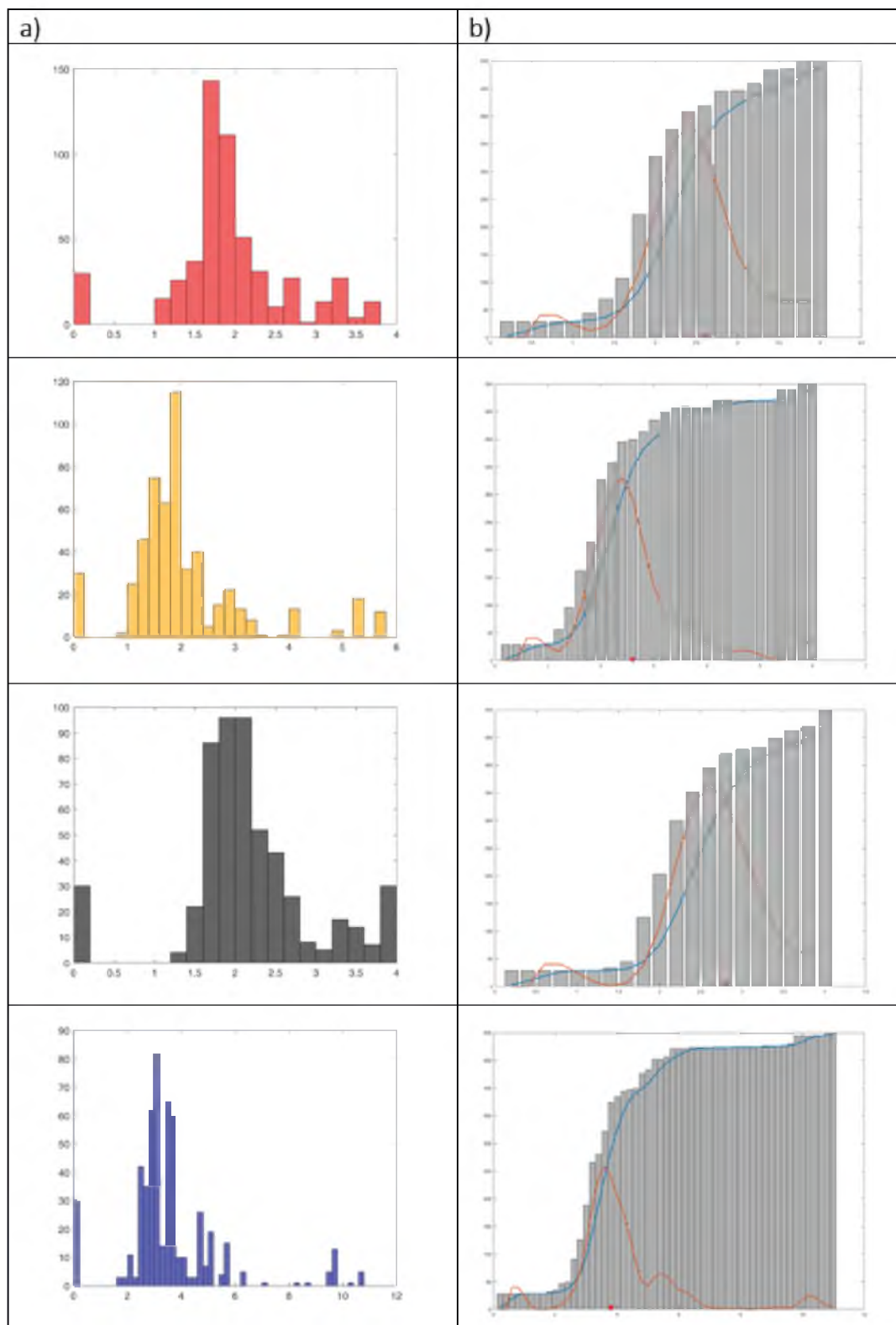
- 1) Przez wyznaczenie progu (położenia linii lokalizującej maksima) na podstawie analizy statystycznej. W tym celu dla każdej z 4 funkcji obliczono odchylenie średnie i standardowe. Próg ustalono na poziomie średniej powiększonej o połowę odchylenia standardowego. Wyniki niestety nie były zadowalające.
- 2) Przez wyznaczenie na podstawie histogramu funkcji lokalnych amplitud poziomu progu, przy przyjęciu założenia, że istnieje minimum histogramu, które oddziela odpowiedzi, w czasie których nastąpiła lub nie nastąpiła reakcja badanego (histogram dwumodalny). Histogramy dla poszczególnych obszarów przedstawiono na ryc. 7a (w lewej kolumnie). Jednak okazało się, że założenie o dwumodalności nie jest spełnione.
- 3) Przez wyliczenie na podstawie standardowego histogramu, histogramu skumulowanego (por. ryc. 7b). Po dwukrotnym „wygładzeniu” (zastosowaniu filtracji uśredniającej) oknem o długości 3 binów histogramu (słupków histogramu) obliczono w oknie o długości 3 binów różnicę (pochodną), a następnie wyznaczono maksimum tej funkcji (punkt przegięcia histogramu skumulowanego). Próg zlokalizowano w kolejnym, względem maksimum, binie histogramu (por. czerwone krzywe i trójkąty na osi odciętych na ryc. 7b).

Za pomocą tej trzeciej metody uzyskano najlepsze rezultaty, zaprezentowane na ryc. 6 (por. grube linie zlokalizowane poniżej odpowiednich funkcji „lokalnych amplitud”). W tabeli 2 zestawiono parametry obliczone dla poszczególnych obszarów.

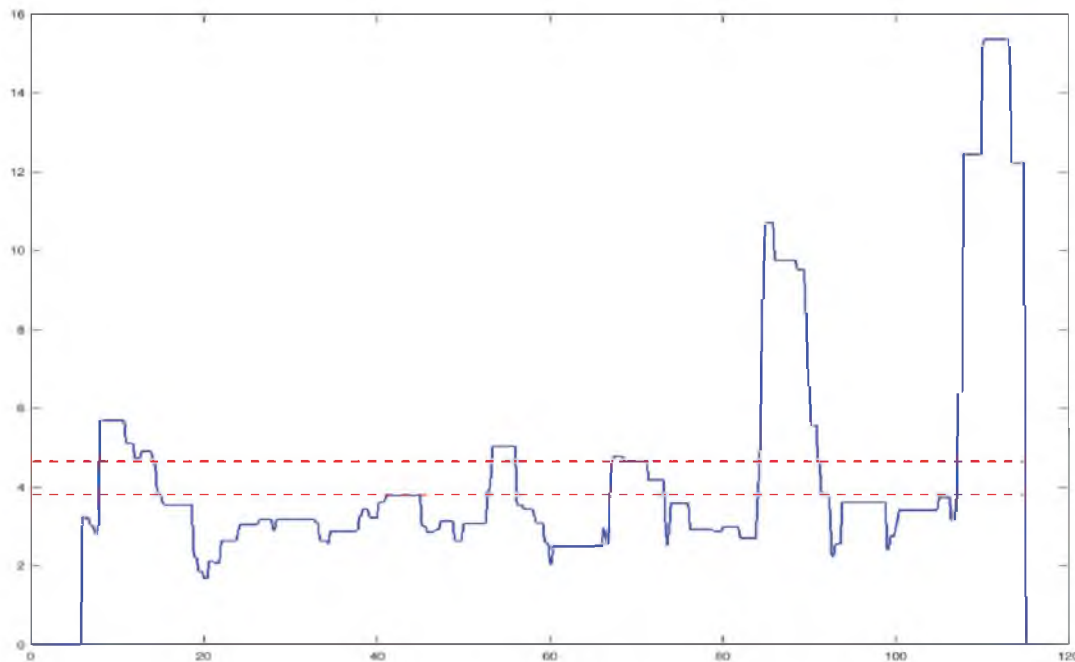
Tabela 2. Parametry obszarów zdefiniowanych w konfiguracji „konf05”

Kolor ramki obszaru	Pole obszaru [piksele]	amplitudy funkcji ‘średnich z obszaru’		amplitudy funkcji ‘lokalnych amplitud’		parametry funkcji ‘lokalnych amplitud’			
		0÷80 [s]	0÷100 [s]	0÷80 [s]	0÷100 [s]	avg	std	próg (met.3)	Δp
czerwony	441	6,3	6,9	3,3	3,7	1,9	0,75	2,6	0,76
pomarańczowy	180	7,3	7,3	4,1	5,6	2,1	1,15	2,6	0,76
czarny	330	6,9	7,7	4,0	4,0	2,2	0,83	2,8	0,7
niebieski	525	8,2	13,6	5,7	10,7	3,6	1,88	3,8	0,87

W 85. sekundzie badania wystąpiły zaburzenia funkcji średnich obliczonych w śledzonych obszarach. Z tego powodu, amplitudy zostały obliczone dla dwóch odcinków czasowych. Średnią (*avg*) i odchylenie standardowe (*std*) obliczono dla zakresu [1, 100]. Próg obliczony metodą trzecią (por. wyżej) nie zależy od interwału czasowego. Interesujący jest parametr Δp wyznaczony manualnie (por. ryc. 8).



Rys. 7. a) Histogramy funkcji 'lokalnych amplitud', b) odpowiadające im histogramy skumulowane, funkcje wygładzone (kolor niebieski), pochodne (kolor czerwony), progi (małe czerwone trójkąci)



Ryc. 8. Zakres zmian prawidłowo różnicującego progu Δp dla obszaru 'niebieskiego'

Jest to przedział zmienności progu, w którym prawidłowo oddzielone są odpowiedzi, w czasie trwania których nastąpiła reakcja, zarejestrowana przez psychogalwanometr, od tych, na które takiej reakcji nie było. Jest on najwyższy dla obszaru „niebieskiego”, który góruje także nad pozostałymi parametrami, zarówno obliczonymi dla funkcji „średnich z obszaru” jak i z funkcji „lokalnych amplitud”. Jak wynika z porównania parametrów, znaczenie ma głównie lokalizacja „okna”. Mniejsze znaczenie ma jego wielkość (powierzchnia), chociaż nie należy definiować zbyt małych okien – por. drugą kolumnę tabeli 2. Poza obszarem „niebieskim”, wielkość odchylenia standardowego nie przekłada się na parametr Δp .

5. Analiza rejestracji (zapisów) uzyskanych podczas wykonywania testu Utah ZCT

Kolejnym etapem badań było przeprowadzenie standardowego diagnostycznego testu poligraficznego, podczas którego badany siedzi w specjalnym fotelu, ma zalecone nieruszanie się i wypowiadanie jedynie krótkich odpowiedzi „tak” lub „nie” na zadawane mu pytania. Badanie dotyczyło sytuacji wyreżyserowanej w warunkach eksperymentalnych (por. relacje z eksperymentu w części II). Przypomnijmy tylko, że osoba badana miała za zadanie zataić w czasie badania, że była na strzelnicy i oddawała strzały do wizerunku kobiety umieszczonego na plakacie.

W czasie badania poligraficznego stosowano test diagnostyczny z rutynowo używanej techniki Utah Zone Comparison Test (Utah ZCT, UZCT) . Test był wykonany w wersji jednozadaniowej⁹. Wykorzystano pytania kontrolne tzw. „kierunkowego kłamstwa” oraz następujące pytania krytyczne:

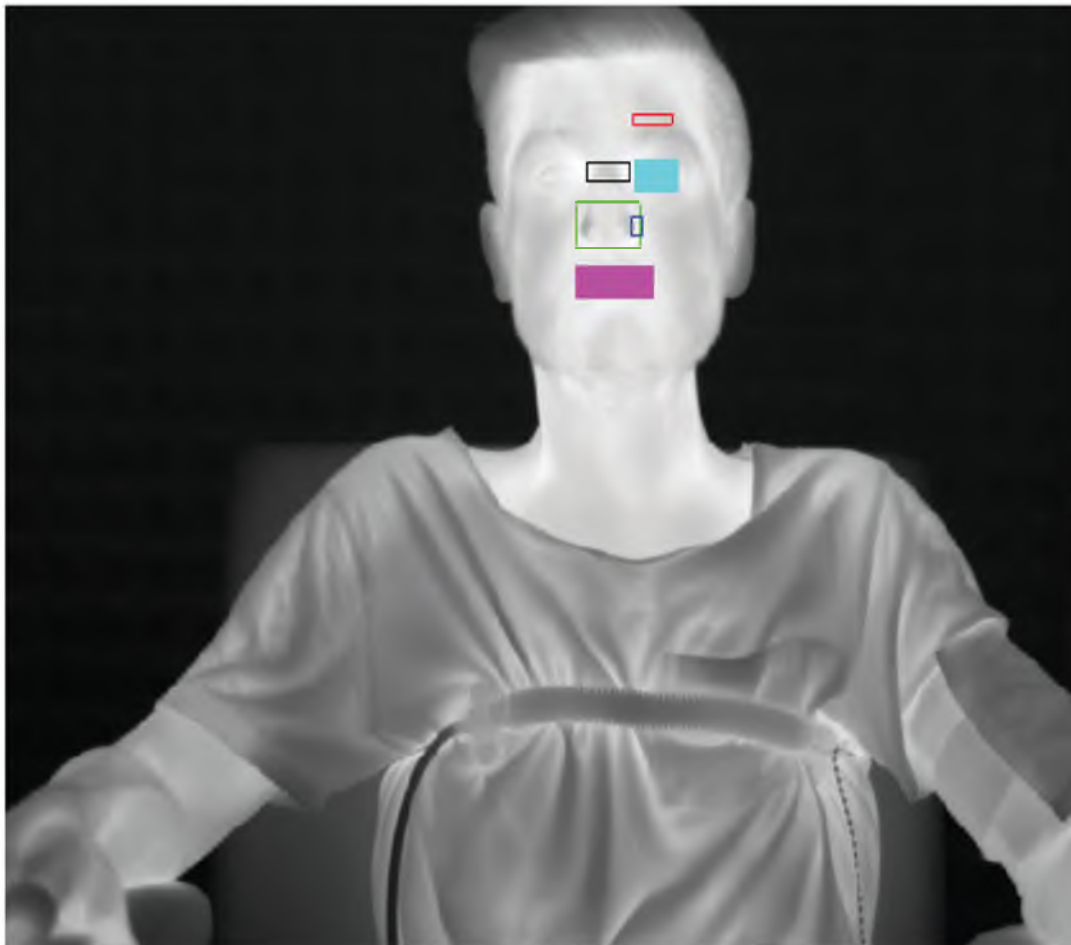
- Czy to ty oddałeś (oddałaś) strzał do wizerunku kobiety?
- Czy miałeś (miałaś) tego dnia bron w ręku?
- Czy tego dnia strzelałeś (strzelałaś) na strzelnicy?

⁹ Technika Utah ZCT została szczegółowo opisana w części II niniejszego tomu.

Dzięki zastosowaniu schematu badania poligraficznego zmiany temperatury twarzy oceniane być mogły przy działaniu wielu bodźców, w postaci pytań testów poligraficznych¹⁰. Badanie składało się z testu stymulacyjnego („z cyfrą”), 3 lub 5 testów diagnostycznych, z których każdy zawierał po 3 pytania krytyczne i kontrolne.

W celu zwiększenia komfortu badanych osób rejestracji filmu dokonano z kamery ustawionej w odległości ok. 2 m od twarzy osoby badanej, a nie jak poprzednio, z odległości ok. 90 cm. Na ryc. 9 pokazano przykładową konfigurację sceny dla tej fazy eksperymentu. Widać założone na ciało osoby badanej czujniki poligrafu i „okna” wyznaczone na obrazie jej twarzy.

/UTAH 5-000025-ramka-125-konf00

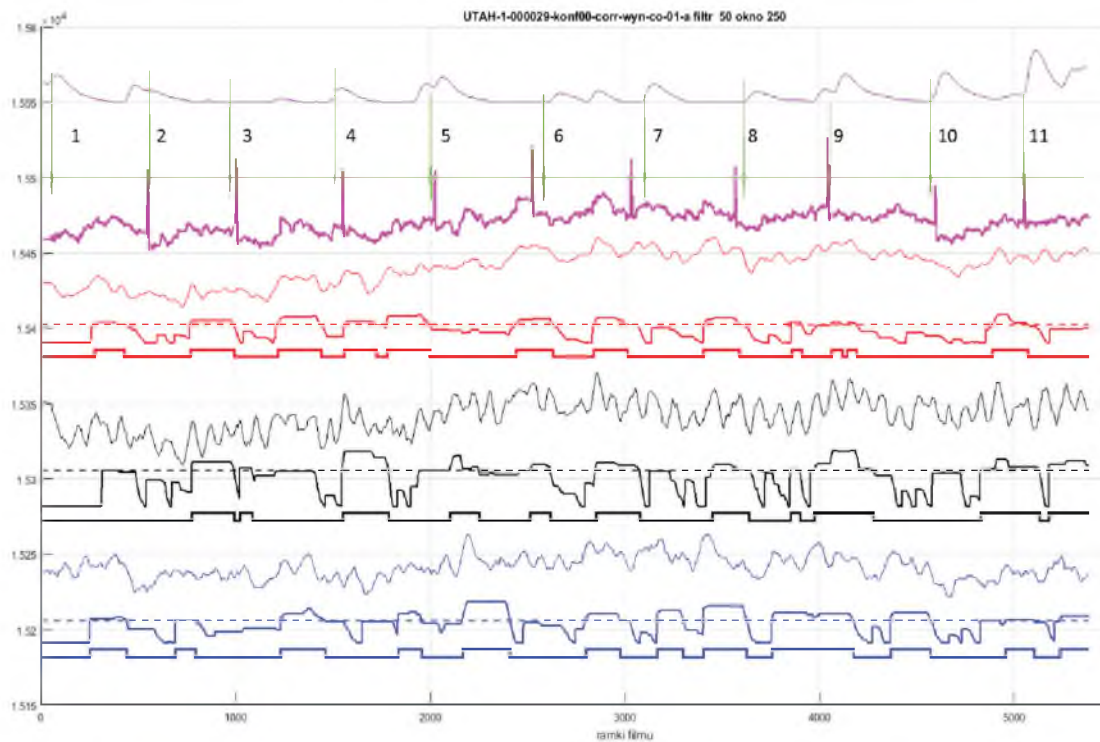


Ryc. 9. Konfiguracja zdefiniowana dla osoby badanej, podobną konfigurację zastosowano w pozostałych eksperymentach

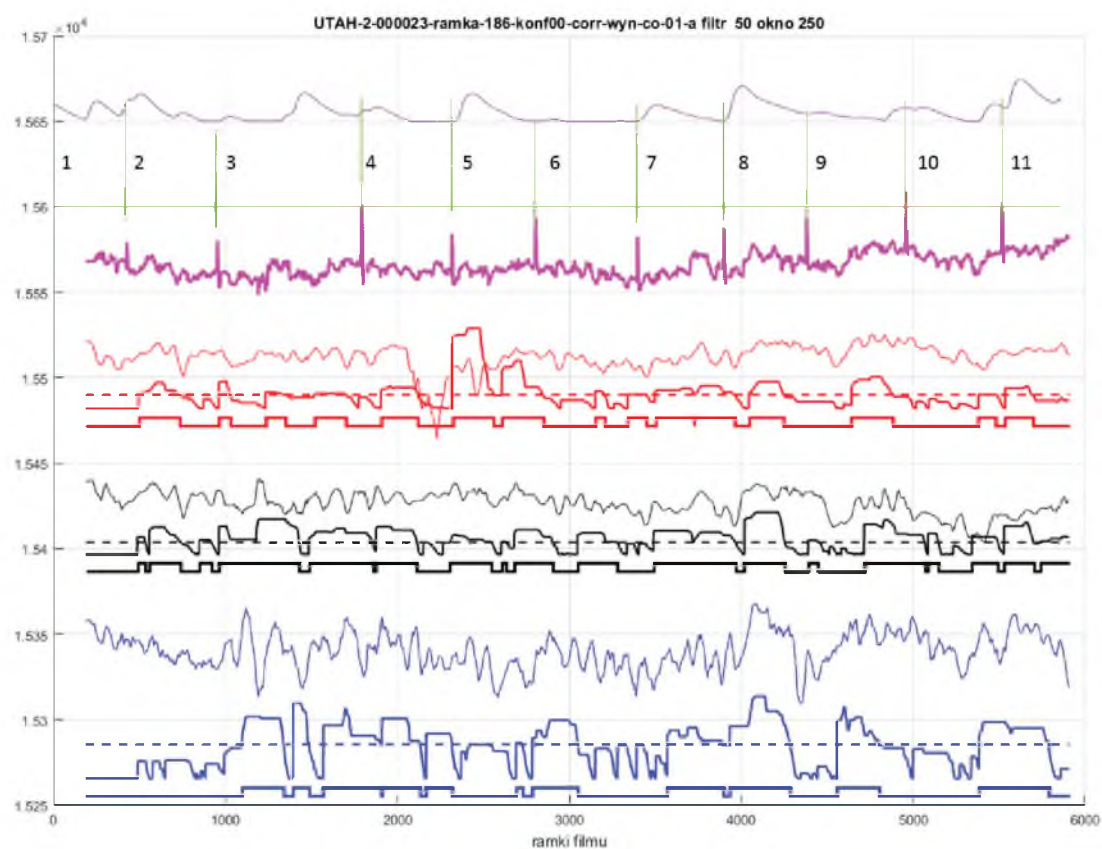
Jak wyraźnie widać na ilustracji, oddalenie kamery termowizyjnej na odległość ok. 2 m od twarzy osoby badanej radykalnie zmniejszyło obraz twarzy, a tym samym powierzchnię „okien” do analizy.

Na ryc. 10, 11, 12, przedstawiono trzy przykładowe zapisy zmian temperatury twarzy rejestrowane równoległe z badaniem poligraficznym. Aby częściowo zrównoważyć zmniejszenie rozmiarów „okien”, analizę przeprowadzono na obrazach o 16 bitowej reprezentacji piksela (por. osie y na ryc. 10, 11, 12).

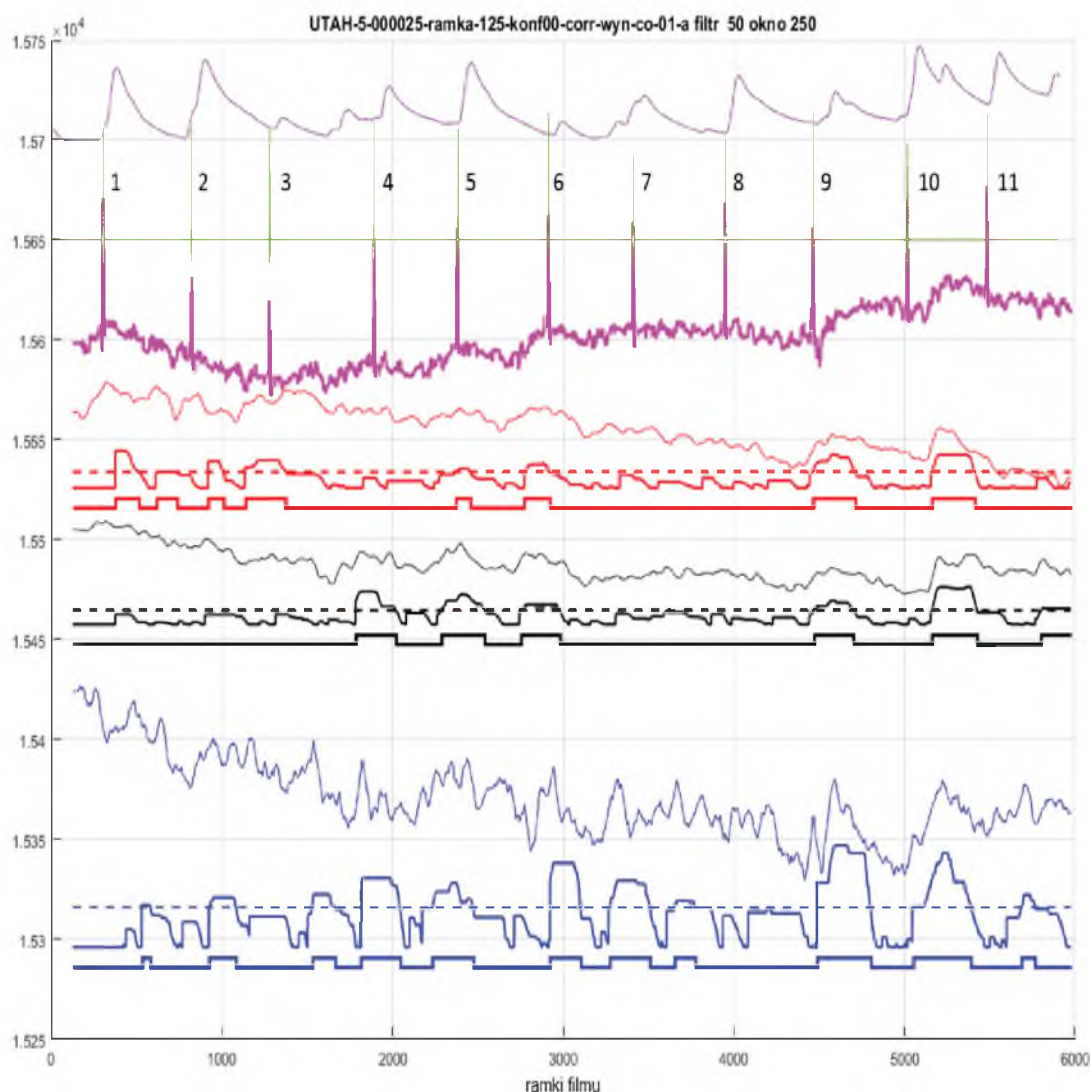
¹⁰ Techniki badania poligraficznego są dokładnie opisane w literaturze. Por. *Kryminalistyka...*, op. cit., s. 426–434.



Ryc. 10. Wyniki analizy filmu z kamery termowizyjnej dla osoby AC144. Kolory zgodnie z konfiguracją pokazaną na rys. 9



Ryc. 11. Wyniki analizy filmu z kamery termowizyjnej dla osoby KT100 seria 2. Kolory zgodnie z konfiguracją pokazaną na rys. 9



Ryc. 12. Wyniki analizy filmu z kamery termowizyjnej dla osoby KT100 seria 5. Kolory zgodne z konfiguracją pokazaną na rys. 9

Progi (położenia linii wykrywających maksima) ustawiano manualnie. Każdy zapis został uzupełniony dwoma wykresami, zlokalizowanymi na górze rysunków (ryc. 10, 11, 12). Kolorem fioletowym oznaczono sygnał z jednego czujnika poligrafu, jakim jest psychogalwanometr (zapis odruchu skórno-galwanicznego). Kolorem zielonym – momenty udzielania odpowiedzi, zarejestrowane przez osobę zadającą pytania. Pomiedzy wykresami umieszczono numery zadanych pytań.

Na podstawie rejestracji poligraficznej i termowizyjnej przedstawionych na ryc. 10, 11, 12 sporządzono zestawienie, które przedstawiono w tabeli 3. Analizę porównawczą, z uwagi na kształt krzywych wykresu, najlepiej jest przeprowadzić, porównując sygnały termowizyjne (wykres zmian temperatury twarzy) z sygnałem z czujnika psychogalwanometru (wykres zmian odruchu skórno-galwanicznego, GSR). Ten ostatni został zaprezentowany na ryc. 10, 11, 12 (wykres pierwszy od góry, w kolorze fioletowym)

Tabela 3. Porównanie wyników detekcji zmian temperatury twarzy i reakcji zarejestrowanych przez poligraf

			pytanie nr										
Probant AC144	Rodzaj badania		INT1	S2	N3	C4	R5	N6	C7	R8	N9	C10	R11
Film .000029 Diagram 1	Poligraf	Pneumo	Rm										
		GSR				Rm	Rm	Rd		Rd			
		CARDIO	Rm			Rd						Rd	
	Termowizja	czoło				x					x		
		oczy		x		x			x		x		
		nos					x		x		x		
Probant KT100	Rodzaj badania		INT1	S2	N3	C4	R5	N6	C7	R8	N9	C10	R11
Film .000023 Diagram 2	Poligraf	Pneumo											
		GSR					Rm			Rd			Rd
		CARDIO		Rm		Rm			Rd				Rm
	Termowizja	czoło			x	x				x			
		oczy				x	x			x			
		nos		x	x					x			x
Probant KT100	Rodzaj badania		INT1	S2	N3	C4	R5	N6	C7	R8	N9	C10	R11
Film .000025 Diagram 3	Poligraf	Pneumo											
		GSR		Rm			Rm		Rd				Rd
		CARDIO		Rm			Rm		Rm				Rd
	Termowizja	czoło			x							x	
		oczy					x					x	
		nos									x	x	

Legenda: x – wystąpienie podniesienia linii temperatury w danym oknie, po rozpoczęciu zadawania pytania;

Rm – wystąpienie reakcji małej, na kanale poligrafu;

Rd – wystąpienie reakcji dużej, na kanale poligrafu.

Na potrzeby tej tabeli postanowiono ocenić globalnie cały poligram (jeden test), i zaznaczyć na nim, gdzie wystąpiły największe reakcje (Rd) oraz dwie kolejne (Rm). Za reakcje na kanałach cardio i GSR uznawano jedynie amplitudę wysokości, nie brano pod uwagę rozbudowania czy długości trwania reakcji. W analizowanych przypadkach zdarzało się, iż reakcje na kanał cardio i GSR były zbliżone wielkością, w takich sytuacjach oznaczano je odpowiednio; jeśli wystąpiły dwie reakcje na jednym kanale podczas testu, to obie oznaczone jako Rd.

6. Podsumowanie

Celem badań było: po pierwsze, znalezienie samodzielnej metody przekształcania obrazu z kamery termowizyjnej w wykres; po drugie, opracowanie metody umożliwiającej ciągłą obserwację i rejestrację zmian temperatury twarzy osoby badanej; po trzecie, znalezienie odpow-

wiedzi na pytanie, czy reakcje emocjonalne człowieka rejestrowane przy pomocy standardowego komputerowego poligrafu (Lafayette LX-4000) mają swoje odpowiedniki w zmianach temperatury twarzy rejestrowanych kamerą termowizyjną (FLIR A655sc).

Udało się wykorzystać i udoskonalić opisaną w literaturze¹¹ metodę przekształcania obrazu z kamery termowizyjnej w wykres.

W celu analizy testowego filmu termowizyjnego nagranych z bliska (odległość ok. 90 cm) wytypowano i zbadano kilka reprezentatywnych obszarów twarzy. Stworzono algorytm wykrywający istotne zmiany temperatury w wytypowanych obszarach. Co więcej, określono obszary, w których analiza daje najlepsze rezultaty. Pełną zgodność uzyskano dla 2 obszarów zlokalizowanych w pobliżu nosa (por. tabela 1, dwa ostatnie wiersze).

Zestawienie zapisów z badań poligraficznych i reakcji termalnych, opisane w ustępie 5, tylko częściowo potwierdzają tezę, że reakcje te są zgodne (równoczesne, równoległe). Wykresy otrzymane w wyniku uśrednienia wartości pikseli z poszczególnych obszarów wydają się zgodne z sobą, natomiast są w niewielkim stopniu zgodne z rejestracjami odruchu skórno-galwanicznego (por. ryc. 10, 11, 12 oraz tabela 3). Nie oznacza to jednak, że nie stwierdzono żadnej zgodności między tymi wskazaniem. Tak na przykład przy pytaniach N9 i C10 (dla filmu 000025) taką zgodność zaobserwowano.

Wydaje się, że przyczyną pogorszenia detekcji termowizyjnej w eksperymencie opisanym w ustępie 5 była zbyt duża odległość kamery od twarzy osoby badanej. Przyczyną takich wyników może też być zbyt mała próba badana. Teoretycznie, przy próbie znacznie większej, taka zbieżność reakcji mogłaby być większa. Dalsze badania muszą po pierwsze być wykonane na większej próbie, a po drugie, kamera termowizyjna musi być umieszczona bliżej twarzy (np. w odległości ok. 90 cm). Eksperymentalnego sprawdzenia wymaga też hipoteza, zgłaszana w literaturze¹², że reakcja skórno-galwaniczna może wyprzedzać w czasie inne reakcje, w tym termiczną.

¹¹ Por. J. Widacki, M. Widacki, J. Antos, *op. cit.*

¹² Por. W. Boucsein, *Electrodermal Activity*, second edition, Springer, New York–Dordrecht–Heidelberg–London 2012, s. 29–36, 130 i nast.

Marcin Gołaszewski, Paweł Zajac, Jan Widacki

Thermal Vision as a Method of Detection of Deception: A Review of Experiences¹

Key Words: thermal vision, detection of deception, instrumental detection of deception, thermal vision and detection of deception

Introduction

Turning pale or red in the face, besides expressive movements (mimic and pantomimic) belonged to the earliest observed symptoms of emotions, which sometimes were even directly treated as symptoms of the lie (Trovillio 1938/1939, Eysenck 1971). Thanks to experimental psychology and physiology, since the end of the 19th and early 20th centuries they have been known to result from changes in the blood supply to the face related to the functioning of the circulatory system, movement of muscles, and chemical changes in the blood and its energetic value (dependent on the amount of oxygen, catecholamines, etc.), a knowledge embracing also the fact that physiological changes accompanying emotions encompass the entire organism and are clearly correlated. Thus, theoretically, it suffices to observe any fragment of the organism to detect emotions, and the number of physiological correlates of emotions, if not unlimited, is certainly very large. Some of them are clearly visible for an external observer (for example, the expressive movements, blanching, blushing, tremors, etc.). Others can be observed only with specialist devices, with polygraph being the best known among them.

The following physiological correlates of emotions are usually selected for observation and examination (see e.g. Hilgard 1972):

- changes in the operation of the heart and the entire circulatory system (blood pressure and heart rate changes, changes of electric phenomena in the muscles of the heart)
- changes in the operation of the respiratory system (changes in the rate of breathing cycles, volume of inspired and expired air)
- changes in electric conductivity (galvanic skin response, GSR)
- changes in muscle tension
- changes in organ volumes (plethysmographic reaction)
- changes in pilomotor reaction
- changes in pupil dilation ('pupillary light reflex')
- body temperature changes.

The last can successfully be observed and registered by a thermal vision camera.

^{*} M. Gołaszewski, P. Zajac, J. Widacki, *Thermal Vision as a Method of Detection of Deception: A Review of Experiences*, "European Polygraph" 2015, Vol. 9, No. 1 (31), p. 5–24.

¹ Project DEC-2013/11/B/H55/03856 funded by National Science Centre.

Thermal vision, also known as thermography, has become one of the most popular and frequently applied methods of imaging heat radiation emitted by physical objects, as it allows remote observation, without the need of intervention into the examined object. Industry, construction, aviation, and medicine are the special areas that can be enumerated within the broad spectre of current practical uses for a thermal vision camera. Thermal imaging is used in medicine for diagnostic reasons wherever a non-invasive measurement of body temperature changes is convenient (Memarian *et al.* 2009; Bauer, Dereń 2014). As a method of observation, measurement, and recording of body (especially face) temperature changes, thermal vision can also be used in experimental psychology, psychophysiology, and their counterparts among forensic sciences. In this last field, its potential use for lie detection is particularly significant.

Thanks to the application of thermal imaging in mid-wavelength infrared (MWIR), a thermal vision camera is capable of detecting temperature changes of specific fragments of the human body, resulting among others from the emergence of emotions that may be related to lying. In this context, thermal vision is an attractive direction in scientific research, as its practical application can be examined in three variants:

- a) instead of a polygraph, independently – when time matters, or at least initial tests must be performed on a mass scale – and/or remotely (contact-free, non-invasively), without the subjects being aware and without the need to employ highly qualified and/or trained personnel (Pavlidis *et al.*, 2002). One of the options to use thermal vision in this way is to set up a system that would collect certain data (e.g. biometric) to allow detection and estimation of the level of threat from various people. It would be possible for example to expand the already operating TOLIS (Traveler On-Line Information System) system, which among others, contains information about the manner of payment for the passenger's ticket: cash or card, with the first one considered more risky, and awarding a greater number of risk points to the traveller (Pavlidis, Levine 2002a)
- b) combined with a polygraph: for possible increasing the precision of (stationary) examinations conducted by adding on top of the traditionally recorded changes (breathing functions, GSR changes, blood circulation changes) those of facial temperature
- c) combined with other technical means capable of remote registration of other physiological correlates of emotions and memory traces (e.g., changes in the voice and the pupillary light reflex) even without the subject being aware.

Symptoms of deception in thermal imaging of the face

Methods that allow high-probability identification of people in whom certain reactions of the organism, attesting to the deception in speech are present are generally known. Such observations may be made particularly through polygraph examination, which typically registers changes in the blood pressure and heart rate, breathing cycle, and electrodermal reactions. Highly encouraging are also the results of experiments conducted by certain researchers and aimed at detection of deception by registering and analysing the temperature of specific fragments of the face under the impact of emotions accompanying lie.

Experiments with a thermal vision camera conducted in the late 1990s brought a discovery that an instantaneous increase of temperature in the periorbital region (see Fig. 1) occurs as a reaction to an external startling stimulus, which results from increased blood flow rate around the eyes (Pavlidis *et al.* 2000; Levine *et al.* 2001). It was established that the discovery can be used for detection of deception, as the blood flow rate accounts for a similar type of physiological data as acquired from the sensors that are part of a traditional polygraph machine (Pavlidis, Levine 2002b).

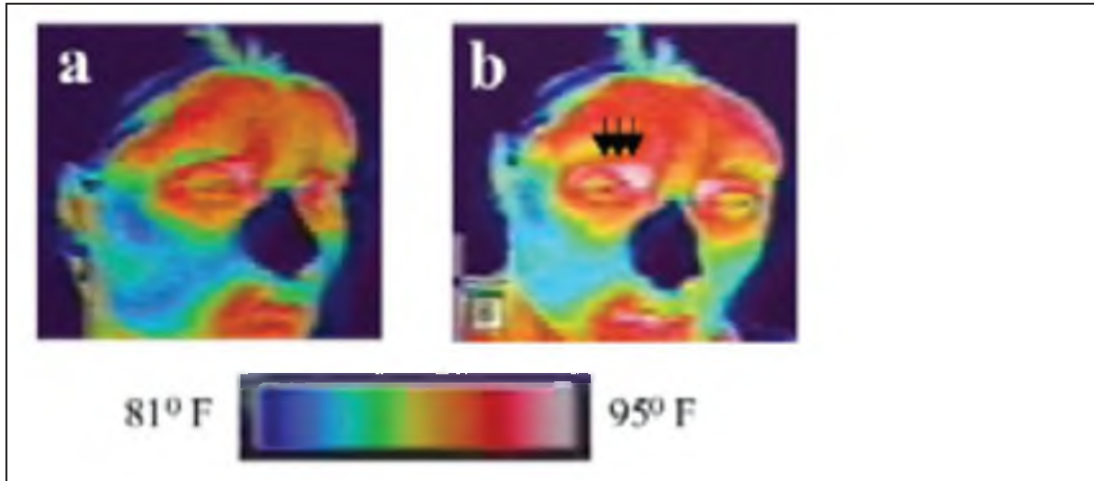


Fig. 1. Thermal images of the face for a subject before (a) and 300 msec after (b) an instantaneous startle. (Arrows indicate local warming in the periorbital region.)

Source: I. Pavlidis, J. Levine, P. Baukol (2000), *Thermal imaging for anxiety detection*, Proceedings 2000 IEEE Workshop on Computer Vision Beyond the Visible Spectrum: Methods and Applications, Hilton Head Island, SC, 16 June 2000, p. 107.

Results of the experimental research conducted by members of the Institute of Electrical and Electronic Engineers (IEEE) show that monitoring of periorbital (P), supraorbital (S) and maxillary (M) regions of the face has a similar diagnostic value as galvanic skin response (GSR; Shastri *et al.* 2009). The phasic component of the periorbital signal is in phase with the GSR signal. The tempo of such changes is different (GSR reactions are more rapid) yet they start at a similar point in time and share a similar trend. Moreover, a phase shift of 180° was observed on the charts between the phasic elements of the curves representing maxillary and periorbital region signals (see: Fig. 2). The reason for this is the link between the signal from the periorbital region and the supply of blood to the orbital muscle which results in increased temperature in the area during stimulation. In turn the signal from the maxillary area depends on the activity of sweat glands, which leads to an inverse effect, that is lowering of temperature.

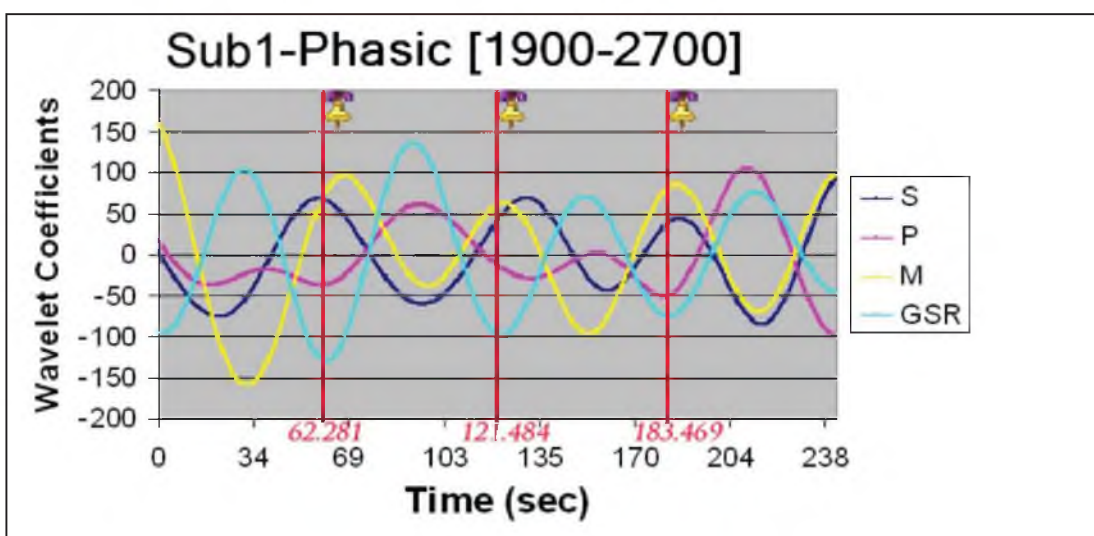


Fig. 2. Phasic components of thermal and GSR signals for subject Sub1.

Source: D. Shastri, A. Merla, P. Tsiamyrtzi, I. Pavlidis (2009), *Imaging Facial Signs of Neurophysiological Responses*, IEEE Transactions on Biomedical Engineering, Vol. 56, No. 2, p. 482.

In an article entitled *The mental nose: thermography and the Pinocchio effect*, E. Salazar Lopez, E. Dominguez, J. de la Fuente, A. Meins, O. Iborra, M.A. Rodriguez-Artacho, and E. Gomez Milan described the potential of using thermal imaging to observe changes of temperature in specific areas of the face as resulting from emotions caused by activities that require mental effort and also linked to lying. The analysis covered readings from the areas around the eyes, corners of the eyes, and the region around the nose. The study group was composed of 20 people. The experiment consisted in performance of three tasks. The task concerning phobias and fears aimed at observation of temperature changes in specific fragments of the face under the impact of emotions emerging while watching images with pictures of drug addicts, spiders, cadavers, and neutral objects. The study group consisted of people suffering from arachnophobia, aichmophobia, necrophobia as well as people suffering from none of the phobias mentioned above. The task related to a mental effort consisted in planning a route in a zoo so as to follow a specific strategy and visit 6 out of 12 places (the results of the Aldeman's Behavioral Assessment of Dysexecutive Syndrome (BADS) were used). In turn, the task related to lie detection was divided into two types of tests (depending on the level of psychological burden). Participants of the examination employing a high level of psychological burden were taken to a laboratory. There, they were informed that the visit was secret and they could not discuss it with anyone. Later, they were given the task to devise an alibi and call a friend or acquaintance. During the telephone conversation, 10 people were to tell the truth, and 10 people were to lie about the course of the day. Participants of the examination with low psychological burden were to provide answers to 12 questions (10 people were to tell the truth and only truth, while the remaining 10 were to provide deceptive answers to 4 questions). The results of the examination concerning lie detection proved that during an exercise requiring a high psychological effort, in 7 in 10 people asked to lie about the course of the day experienced an increase of temperature around the nose by 1.3°C and by 0.9°C in the region of the forehead was recorded while they were planning the conversation, i.e. building their alibis. In turn, in the case of lying during the telephone conversation, a drop of temperature by 1.1°C was recorded around the nose in 80% of participants. To compare, no changes in face temperature were recorded at all in the group telling the truth (Salazar *et al.* 2012).

I. Pavlidis and J. Levine transformed the raw thermal data into data concerning the blood flow rate by using a thermodynamic model that assumes that the blood flow rate is conversely proportional to the square of skin temperature deviation from core temperature of the human body. An stumbling block in this model is the impact of propagating heat generated by post-prandial metabolism, which provides a disruptive factor: the reason why the subject should fast for a number of hours before the examination. Cooperating with psychologists from the US Department of Defence Polygraph Institute (DoDPI), the two scientists conducted an experiment with 32 subjects. Some of them were to perform a mock crime (jabbing a mannequin with a screwdriver and stealing \$20), while the others had no connection to the event. Subsequently the whole group were subjected to a polygraph examination using federal ZCT and GKT tests. Pavlidis and Levine's method consisted of three stages: image acquisition, physiological correlation (calculation of blood flow rate), and standard classification – final opinions concerning truthfulness of the subjects' answers (Pavlidis, Levine 2001). Images of the subjects' faces were recorded while they were answering questions asked during a polygraph examination. A cooled mid-infrared camera operating in mid-wavelength infrared, with sensitivity of 0.0 25°C was used. The temperature scope was set from 29°C to 38°C which corresponds to the external temperature of human face. To eliminate the impact of lighting on the measurements, the examination was conducted in a darkened room. In the case of the ZCT test, the average length of recording was 300 frames, at the speed of 30 frames per second (fps), and for the GKT was half as short. Two regions of the face taken into account in the analysis – the area around the eyes and the forehead of the subject – were defined on the first recorded frame. This followed the assumption that the object of the study is relatively stationary, as in this case movement is a disruptive factor. After the exclusion of materials that could not be appropriately assessed for a variety of reasons, data of 18 subjects were considered analysis-worthy, with most adequate being the answers to the question no. 10 from the federal ZCT sequence. The 'forehead' signals proved insufficient to distinguish truthful from deceptive subjects. The case was otherwise with the signals from the area around the eyes. They provided the grounds

to state that the reaction to the question occurs in two stages. While the question is being answered, the curve defining the blood flow rate in time rises slightly in the case of all the subjects. Not much changes in the second phase if the subject answers truthfully. The difference is, however, visible in the case of people providing deceptive answers: in their case the curve grows rapidly, as an expression of increased blood circulation around the eyes under the impact of anxiety. The authors explain that this data is perfect for a polygraph examination, because (like in the case of the polygraph), the interesting element is the relative speed of the rising of the curves (see Fig. 3).

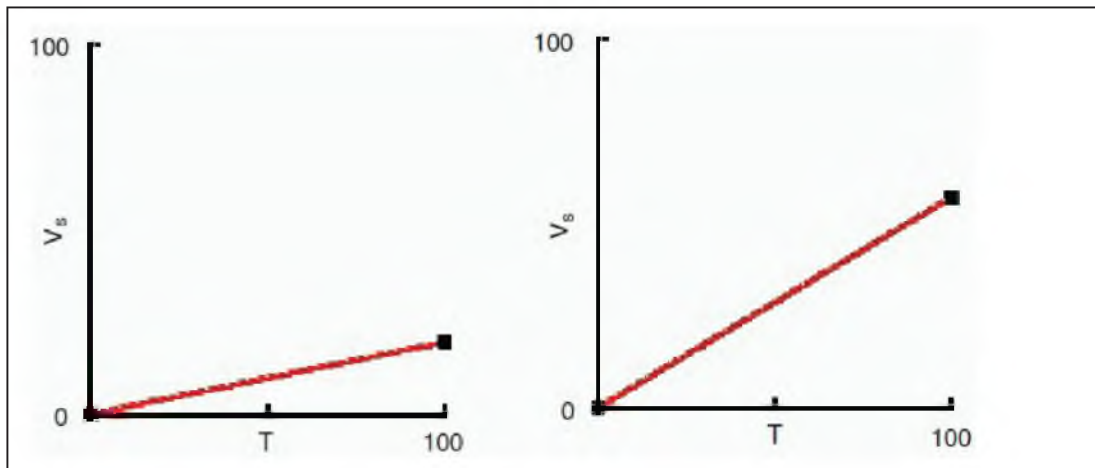


Fig. 3. Normalised lines characterising the slopes of the average periorbital blood flow rate in time T . The slope of the answer session appears quite steep.

Source: I. Pavlidis, J. Levine (2002b), *Thermal Image Analysis for Polygraph Testing*, IEEE Engineering in Medicine and Biology, Vol. 21, No. 6, p. 61.

The definition of threshold values makes use of the so-called Otsu algorithm. If the slope product of the curve was lower than the threshold, the subject was classified as truthful, was greater – deception was concluded. The value of deviation was the angle between the curve and the horizontal line of the chart defining time following the answer (151 frames). The threshold used for the binary honest/deceptive classification was set at 67° . Finally, the proportion of correct decisions while analysing facial temperature changes amounted to 84%, and was insignificantly different from the correctness level obtained in the experiment through the traditional method of polygraph chart analysis (78%). Pavlidis and Levine recognised that further studies should among others focus on the improvement of the assumed model of heat propagation by estimation of the factor related to the heat produced through metabolism, design of an algorithm to allow tracing of head movements during the test, definition of a mesh of specific critical points on the face where blood flow rate will be monitored.

An experiment concerning the possibility of combining thermal imaging with a polygraph examination to detect deception was conducted by D.A. Pollina and A.H. Ryan of DoDPI. To test the hypothesis that temperatures of specific areas in the face change while answering relevant and comparison questions related to deception, 14 regions of the face were selected for analysis. It comprised readings from around the mouth, ears, neck, and eyes, with the maximum values of temperature recorded and collected bilaterally (e.g. from both the ears). The average from the pair of readings was used for statistical analysis. The study group consisted of 16 people. Before starting the test, participants were divided into two subgroups. The first subgroup of ‘the innocent’ was told that – because they were ‘innocent’ – they were to answer questions truthfully. In turn, the task of the other subgroup was to commit a mock crime, and later to lie about it during a polygraph examination, making use of a Zone Comparison Test (ZCT) with two repetitions. A recording with an infrared camera (256×256 FPA, sensitivity $<0.1^\circ\text{C}$) began at the moment when relevant and comparison questions were asked. A model of binary logistic regression was used in the examination. The following correlations were

tested: thermal imaging – breath, thermal imaging – GSR, and thermal imaging – blood pressure. The highest correlation was proved for the thermal imaging – blood pressure pair. The results of the test proved that the regions around the nose and eyes demonstrate highest temperature fluctuations connected directly to lying. The authors believe that these regions can be considered the relatively most efficient deception predictors, and the most precise overall results were acquired in combination with a traditional polygraph examination (Pollina, Ryan 2004).

Table 1. Results of the experiment conducted, denoting the number of correct indications (hits) while using polygraph, and combined examination: polygraph with thermal imaging of the region around eyes and nose.

Cutoff Score	Polygraph Measures			Polygraph & SST: Eye			Polygraph & SST: Nose			Polygraph & SST: Eye, Nose		
	Hit	Miss	Inc.	Hit	Miss	Inc.	Hit	Miss	Inc.	Hit	Miss	Inc.
.5/.5	16	8	0	19	5	0	21	3	0	20	4	0
.6/.4	16	4	4	17	2	5	16	3	5	18	3	3
.7/.3	15	2	7	15	2	7	14	1	9	16	2	6
.8/.2	14	0	10	12	2	10	12	1	11	15	0	9
.9/.1	7	0	17	11	0	13	12	0	12	12	0	12

Source: D.A. Pollina, A.H. Ryan (2004), *The relationship between facial skin surface temperature reactivity and traditional polygraph measures used in the psychophysiological detection of deception: a preliminary investigation*, Polygraph, Vol. 33 (3), p. 150.

The DoDPI scientists conducted also two other experiments with thermal vision (Pollina *et al.* 2006). In both cases, infrared cameras were used during a CIT (Concealed Information Test) polygraph examination. Participating in the first experiment were 30 people, of which number records from the examination of 15 participants could be used finally. The appropriate frame with the beginning of the verbal answer of the subject was selected visually and included into the analysis of three frames distant 0.33 s from one another, and preceding and following the given answer. For each recorded frame, the maximum and minimum temperature of individual pixels was recorded for the analysed timeframe. It was discovered that differences in skin surface temperature in people providing truthful and deceptive answers depend on time segments of the recording and the type of test stimulus. This in turn proves that the reactions of a subject are related to the processing of information concerning a specific stimulus (question) and preparation to the answer. Interestingly, statistically significant measurements were acquired only in the case of the right half of the face. This can be explained by emotional reactions which the right hemisphere of the brain is responsible for or by the nervous activity conditioned by the sympathetic part of the autonomous nervous system aimed at immediate supply of oxygenated blood to the eyes. In the second experiment, procedures from the first were replicated but on a slightly larger sample (39 subjects with data from 24 examinations qualified to the final analysis), a different camera was used, and the system for tracing head movements was more sophisticated. Thanks to this, 91.7% of the subjects were correctly classified for truthfulness/deception.

A similar experiment making use of the CIT procedure and detection of deception in utterances (false declarations concerning lack of knowledge about details of an event) in participants who perpetrated a mock crime (stealing a wallet) was conducted by K.K. Park, H.W. Suk, H. Hwang, and J-H. Lee. They analysed the readings from periorbital regions of the face (a triangle-shaped area between the eye and the nose on both sides of the nose, and also the area around the eyes). The study group was composed of 34 people. The examination made use of thermal imaging combined with analysis of variance (ANOVA). The participants were asked questions in accordance with the CIT procedure. According to this assumption, the person who stole the wallet will have critical knowledge of the subject. In the case of 'guilty' people answering irrelevant questions, a temperature of specific regions of the face recorded within 10 seconds from asking the question was by 0.55°C lower than in the case of answers provided to significant questions. In turn, in the 'innocent' participants, whose only task

was to send an email from a laboratory, temperatures of specific fragments of the face while answering the significant questions were by 0.33°C lower as compared to irrelevant questions (Park *et al.* 2013).

Studies making use of an integrated measurement with the use of an infrared camera and polygraph were also conducted by K. Kim and J-H. Lee of Chung-Ang University in South Korea. The experiment made use of CQT (Comparison Questions Test). The study group were 37 people. Participants were divided into two subgroups. The first of them were supposed to steal a wallet from a computer lab and hide the evidence. The only task given to participants from the other subgroup was to send an email from the laboratory. Polygraph examination and examination making use of the infrared camera were made at the same time to develop and assess an integrated technique for observation and recording of physiological changes accompanying deception. Results analysis made it possible for the researchers to identify correctly 73.3% of participants in the study who answered deceptively the relevant questions (Kim, Lee 2014)

Are the changes in temperature of specific regions occurring in the face while lying related to brain activity (EEG)? The answer to this and other questions through thermal imaging was sought by U. Jain, B. Tan, and Q.P. Li. The scientists conducted an experiment in which the cohort of subjects was composed of 16 people. Some of them committed a mock crime (stealing a ring), and later all the subjects were to answer questions related to the circumstances of the committed theft. Before the start of the test, participants were divided into two subgroups. Every other one went to an electroencephalographic examination combined with thermal imaging. The other half of the group were asked questions in line with the CIT procedure (i.e. they were asked about details related with the mock crime). The analysis covered readings from the internal and external regions around the corners of the eyes, fragments of the nose, the tip of the nose, and the external corners of the mouth. Using the infrared, the researchers reached 83% of correct indications of deception. The error level was estimated at max. 0.485, and min. 0.165, which means that the lowest level of correct lie detection amounted to 51.5%, and the highest to 83.5% (with 20 repetitions). In turn, EEG made it possible to achieve 90% of correct indications of deception in utterances (Jain *et al.* 2012).

An approach to thermography different than that of majority of researchers was demonstrated by P. Tsiamyrtzis, J. Dowdall, D. Shastri, I. Pavlidis, M.G. Frank, and P. Ekman. They recognised that what is needed is a solution that – with instrumental and automated detection of deception – would make it possible to develop conditions as close to real life as possible and ensured a statistically significant number of participants (a comparison of the fundamental assumptions present in various experiments with thermal imaging in nearly the last two decades is presented in Table 2). To achieve this, 39 people were randomly drawn from a population of 250 selected by M.G. Frank and P. Ekman. Each of them filled in a questionnaire involving biographic and personality questions. Later, the participants were grouped into twos, composed of people with opposing beliefs and views of the world. One person in each pair (the future subject) was informed that there is an envelope with a cheque in the hall. The task of the person was to find the envelope and to make a decision concerning its potential theft. The cheque was made to a group that the finder certainly did not support, which was to encourage the theft. Independent whether the theft occurred or not, the intended recipient of the cheque (or a supporter of the organisation that was to receive it) interrogated the subject. To strengthen the motivation of the experiment further, envisaged for the participants in the experiment were rewards (pecuniary) and punishments (in the form of money and doses of noise through headsets) were envisaged. Depending on whether one managed to convince the interrogator about one's innocence or not, and also on the correctness of the interrogator's opinion, various options of punishments and rewards for the people in a given pair and their groups were present. The questions used in the interrogations were designed by a team composed of experts in homicide, counterterrorism, and behaviourists. Most of the questions aimed at obtaining maximum number of non-verbal behaviours useful for analysing (Tsiamyrtzis *et al.* 2007). The questions were as follows:

1. Describe in detail everything you saw down the hall.
2. Describe your actions while you were down the hall.
3. What was going through your mind regarding whether to take or leave the cheque?

4. Did you remove the cheque from the envelope?
5. Is there anything else you wish to tell me about the cheque or what you did?
6. Is everything you have told me about the cheque the truth?
7. What would you say if later I determined you lied to me about the cheque?
8. Where did you put the cheque?
9. What should happen to a person who took a cheque like this?
10. Have you ever told a lie to get out of serious trouble?
11. I know who that cheque was made out to. It is very understandable that you would not want to see money going to that organization. If I were in your shoes, I probably would have do anything in my power to avoid having money go to this organization. Now tell me the truth, did you take that cheque or not?
12. If you were to take a polygraph (lie detector machine) test, what do you think the results of this test would be?
13. Why do you think that someone would take this cheque?

The subjects provided free answers to the questions above, only once, and the interview was run at a natural speed. Later, the data recorded with an infrared camera were analysed. The top 10% of the hottest pixels in the periorbital region were selected and had average temperature, corresponding to the mean temperature of the vascular system in the internal corners of the eyes, calculated; and the recorded signal was filtered with a low-pass filter. The researchers noticed that various people react with different intensity to the same stressor. That is why they compared fluctuations of the temperature signal with the data from the entire interview with the subject in their calculations. If the deviation in the question-and-answer section area was greater than during the entire interview, the answer was considered deceptive, and if it was smaller – an opposite conclusion was drawn. Results close to 0 could be considered inconclusive. Altogether, a level of 87.2% correct indications was reached.

D. Shastri, P. Tsiamyrtzis, and I. Pavlidis noticed that the heuristic method of selecting 10% of highest temperature pixels is strongly dependent on the qualifications of the camera operator and selection of the preliminary region of interest (Shastri *et al.* 2008). Instead, they presented a new method for precise distinction of the signal from the periorbital region. In the first stage, the area is segmented, and in the second – it is tracked frame by frame, while the third contains the refinement of the heat signal from unnecessary interferences. Segmentation is effected through the choice of two points situated in the internal corners of the eyes in the starting frame, and localisation of pixels with highest temperature, which are later used in an adaptive fuzzy connectedness algorithm. It was assumed that if the physiological change accompanying the critical question will be greater than the change in the base line, the subject's answer will be considered deceptive. Otherwise, the subject is classified as telling the truth. In experiments involving 39 subjects, deceptive answers were correctly discovered in 82% of the cases.

In their experiments most researchers focused on the periorbital region of the face, yet there were also other attempts, for example, thermal imaging of the nose region. The neurophysical reactions caused by stress that appear in the region are manifested through short-term increased sweating (Shastri *et al.* 2012). In turn, Z. Zhu, P. Tsiamyrtzis, and I. Pavlidis discovered that the muscle responsible for wrinkling the forehead is more active than usually when a subject experiences persistent stress (Zhu *et al.* 2007). The signal from the periorbital region provides information concerning momentary stress related to alarming stimuli, while the signal from the supraorbital region accounts for information concerning long-term stress resulting from mental involvement. These allowed to put up a hypothesis that deceptive individuals can be more strongly involved mentally, because they make up various stories. Thirty people were invited to the test. It was discovered the temperature changes during the interrogation result from two phenomena: the interrogation itself, and lying. In the case of deceptive individuals, the two factors become superimposed. The rate of correct detection of deception based on facial temperature changes around the forehead amounted to 76.3%.

Table 2. Thermal vision in detection of deception: a summary of selected experiments

Authors	Type of subject behaviour in the experiment	Region of interest (ROI) in the face	Diagnostic value of the method applied
I. Pavlidis, J. Levine, 2001	The subjects supposed to lie in the test were to jab a mannequin with a screwdriver and steal money (\$20). The task of the innocent subjects, unconnected to the event, was to answer truthfully.	the forehead, and primarily the periorbital region	78%
D.A. Pollina <i>et al.</i> , 2004; D.A. Pollina <i>et al.</i> , 2006	The task of the subjects supposed to lie in the test was to jab a mannequin with a screwdriver and steal money. The task of the innocent subjects was to answer truthfully.	14 regions in the face, around the mouth, eyes, nose, and neck; mainly the periorbital region	91.7%
P. Tsiamyrtzis <i>et al.</i> , 2006	Discovery of an envelope with a cheque, its optional theft, and later participation in an interrogation. The selection of the potential thief and interviewer (beneficiary of the cheque) was based on opposing traits of personality and different views of the world. Moreover, an extended system of motivation was used.	periorbital region	87.2%
Z. Zhu <i>et al.</i> , 2007	no data	the forehead (supraorbital region)	76.3%
D. Shastri <i>et al.</i> , 2008	no data	periorbital region (internal corners of the eyes)	82%
L. Warmelink <i>et al.</i> , 2010.	The subjects invited to experiments in the airport were to answer truthfully about the destination of their travel but had to lie about the reasons why they travel.	no data	64-69%
E. Salazar Lopez <i>et al.</i> , 2012	The subjects were given the task of making a fictional alibi and lie about the course of the day in a telephone interview.	the area around the nose, corners of the eyes	
U. Jain <i>et al.</i> , 2012	The subjects were to steal a ring and later lie about the mock crime.	corners of the eyes (internal and external), nostrils (external edges), nose and corners of the mouth	83.5%
K.K. Park <i>et al.</i> , 2013	The subjects who were to lie in the test were to steal a wallet. The task of the subjects who were to tell the truth was to send an email from the lab.	the triangle-shaped area between the eye and the nose, on both sides of the nose; also around the eyes (principal)	98.89%
K. Kim <i>et al.</i> , 2014	The subjects who were to lie in the test were to steal a wallet. The task of the subjects who were to tell the truth was to send an email from the lab.	no data	70.6% – 73.3%
B.A. Rajoub, R. Zwiggelaar, 2014.	The subjects were to learn a false personal profile with data about their education, family, and job. Subsequently, they were supposed to lie during an interview about who they are and what they do in their lives.	corners of the eyes (by the nose) – 17x17 pixels.	87%

Source: own materials or the authors

Polish experience

Although the theoretical possibility of using thermal vision cameras in detection of deception was recognised in Poland already in the 1970s (Kołecki 1979), the subject was discontinued for many years. Only recently, the interest in the method increased, which resulted in the first experimental attempts at using an infrared camera for that purpose. The experiments conducted so far (initial investigations by Polakowski, Kastek, Pilski 2011; Staszal, Wojtarowicz, Zając 2013), have not gone beyond the general support for a well-known fact that an infrared camera can record facial temperature changes as physiological correlates of emotions. Moreover, they validated problems resulting from the need to stabilise the head of the subject for the time of the examination put forth by other authors.

Problems in infrared studies conducted so far

Some researchers emphasise that thermal changes in specific regions of the face result not only from the subject being deceptive, but may correlate with the activity of specific groups of facial muscles, metabolism, sicknesses, and ambient heat emission (Pollina *et al.* 2004; Warmelink *et al.* 2010; Rajoub *et al.* 2014). Others point to the need for standardisation of the research method, technique, and high resolution equipment. It is necessary to develop more advanced methods (algorithms) for tracking the face, and also to develop a simpler, faster, and more precise way of conducting analyses of the collected thermal data (Park *et al.* 2013). Moreover, a more decided approach to the statistical analysis of maximum thermal values is emphasised: the analysis would focus not on a single pixel but on a pixel array from thermal images (Park *et al.* 2013).

Depending on the region of the face which is our region of interest, seemingly simple obstacles emerge and render precise measurements practically impossible. Such obstacles include e.g. spectacles, as glass may conduct neither far- nor mid-wavelength infrared, in the consequence distorting the records concerning temperature changes around the eyes. A long and thick fringe in turn interferes with measurements of forehead temperature.

A major problem while recording changes taking place in specific regions of the face are also the movements of the subject's head. It is necessary to limit the movements of the subject to minimum while performing point imaging. Steps aimed at limiting the instabilities in data recording caused by the subject's movement were taken. An experiment aimed at testing system efficiency – a network for tracing temperature changes in specific fragments of the face – was conducted among others by D. Shastri, P. Tsiamyrtzis, and I. Pavlidis. Their study proved that a network system for tracing temperature changes provides highly constant efficiency (20 trials), with an average precision error of one pixel. Efficiency of monitoring network is decidedly greater than that of a single tracker while the subject is moving the head. To optimise the interaction of each individual tracker, John von Neumann's Game Theory was used. The adaptation of the theory aimed at observation of individual trackers as players in a game, whose goal was to increase the impact of an individual tracker by building coalitions with other trackers to achieve a common goal. The results of the study proved that a system for network tracking has a greater efficiency than a single tracker (Shastri *et al.* 2008).

The goal behind the creation of a more perfect system for tracking temperature changes of specific regions of the face was set up by P. Tsiamyrtzis, J. Dowdall, D. Shastri, I. Pavlidis, M.G. Frank, and P. Ekman. The need was justified by the fact that even a momentary loss of the tracking path results in sudden leaps in the recorded temperature signal, which in such a case no longer reflects true reactions (Tsiamyrtzis *et al.* 2007). The scientists designed a so-called tandem tracking system, which is expected among others to be more resilient to head movements of the subject. It is a modified system for tracking a single (periorbital) region using information from another (central) region. A technique for adjusting the pattern based on the selection of the object of interest in the initial frame and later finding regions most akin to the original pattern in the following frames was used. This innovation helps to reduce errors in face tracking construed as any diversion of the point of tracking from the

periorbital region for a period extending one frame. In this way, the recording of temperature changes in time became more adequate and legible (see: Fig. 4). The tracking system retains capacity to recover in the cases of occlusion incidents and turns up good results (accuracy of up to ± 2 pixels compared to the actual target), the translational speed was ≤ 3 pixels per frame, and rotational speed was $\omega \leq 5$ radians per frame. This luckily corresponded to the actual subject dynamics.

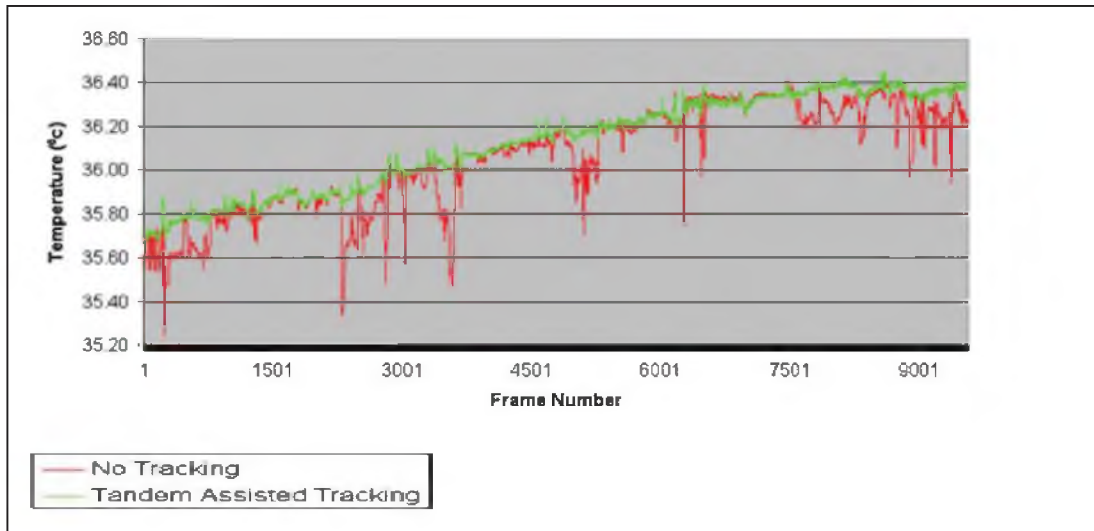


Fig. 4. A comparison of obtained temperature signals from the periorbital region of a subject (without tracking and with tandem tracking) in time.

Source: P. Tsiamyrtzis P., J. Dowdall, D. Shastri, I. Pavlidis, M.G. Frank, P. Ekman (2007), *Imaging facial physiology for the detection of deceit*, International Journal of Computer Vision, Vol. 71, p. 199.

As far as the technological aspects for detection of deception through combining a polygraph examination with thermography are concerned, computer software processing all the recorded data in real-time and making it presentable in a single polygraph chart would be handy.

Despite pointing to the still existing limitations of various kind that impact the precision of the method and its practical applicability, most researchers who participated in the thermal vision experiments perceive a vast potential in the method. Not all, however, approach this type of examinations with enthusiasm, which is especially true of the use of an infrared camera as a tool for routine control of integrity. For example, a group of experts specialising mostly in non-instrumental methods of deception detection (including A. Vrij) believe that thermal imaging will lead to an excessive number of false cases of detection of deception, because there are various sources of anxiety, and no significant changes between the truthful and lying individuals are observed before embarking on the interview (Warmelink *et al.* 2010). As far as an infrared camera can be taken considered in mass security checks, for example, in airports, this makes sense only during an interview and not remotely, and even more so without making the subject aware. Moreover, in the experiments they conducted, people running interviews who did not account for the recording from infrared cameras formed more accurate opinions. That is why researchers suggest that it is better to invest funds into the enhancement of interviewing methods rather than into upgrading thermal imaging equipment.

Conclusions

It seems to go without saying that the method for thermal registration of facial temperature changes as a correlate of emotions is interesting for psychology, psychophysiology, and neuropsychology (Milan *et al.* 2012). Moreover it becomes highly promising in the detection of deception and hidden information (Jain *et al.* 2012).

As the experiments conducted so far show, the best results are obtained when the method is combined with classical polygraph examinations (Pollina *et al.* 2004; Kim *et al.* 2014). At the current stage of research, this is however not yet a solution that could be used efficiently as an independent method for detecting deception. Besides minor obstacles of technical nature, which can be eliminated gradually in the successive experimental studies, application of this method for the detection of deception has as yet not been possible in practice, especially on a mass scale, without the awareness and cooperation on behalf of the subject. Secondly, an examination of facial temperature changes as an independent method of deception detection would be based on the assessment of only a single physiological coordinate of emotions (moreover: being secondary to the changes in the circulatory system and muscle tension), while a classical polygraph examination is based on a simultaneous assessment of at least three such correlates.

The future of this method of lie detection may be perceived in a) treating it as supplementation of a classical polygraph examination providing another emotion coefficient subjected to assessment (facial temperature changes – mostly in periorbital region); b) using it in conjunction with other methods that allow contact-free (i.e. remote) observation of at least two other physiological indicators (emotion correlates) – e.g. changes in the voice, eyeball movements, etc. This, however, is a task by far more difficult.

References

- Bauer J., Dereń E. (2014), *Standaryzacja badań termograficznych w medycynie i fizykoterapii*, Acta Bio-Optica et Informatica Medica Inżynieria Biomedyczna, Vol. 20, No. 1, p. 12.
- Eysenck H.J. (1971), *Sens i nonsens w psychologii*, Warszawa, pp. 87–88.
- Hilgard E.R. (1972), *Wprowadzenie do psychologii*, Warszawa, pp. 243–245.
- Jain U., Tan B., Li Q. (2012), *Concealed Knowledge Identification using facial thermal imaging*, Li Creative Technologies, Florham Park, NJ, USA.
- Kim K., Lee J-H. (2014), *Detecting deception using integrated thermal imaging and polygraph*, International Journal of Psychophysiology, Vol. 94.
- Kołecki H. (1979), *Kryminalistyczne zastosowania termografii*, Warszawa.
- Levine J., Pavlidis I., Cooper M. (2001), *The face of fear*, The Lancet, Vol. 357, No. 9270.
- Memarian N., Venetsanopoulos A.N., Chau T. (2009), *Infrared thermography as an access pathway for individuals with severe motor impairments* [online], Journal of Neuroengineering and Rehabilitation, Vol. 6, No. 11, [accessed on: 27.03.2015], www.jneuroengrehab.com/content/pdf/1743-0003-6-11.pdf.
- Park K.K., Won Suk H., Hwang H., Lee J-H. (2013), *A functional analysis of deception detection of a mock crime using infrared thermal imaging and the Concealed Information Test*, Frontiers in Human Neuroscience, Vol. 7.
- Pavlidis I., Eberhardt N.L., Levine J.A. (2002), *Seeing through the face of deception*, Nature, Vol. 415, p. 35.
- Pavlidis I., Levine J.A. (2001), *Monitoring of periorbital blood flow rate through thermal image analysis and its application to polygraph testing* [online], Proceedings of the 23rd Annual EMBS International Conference, 25–28.10.2001, Istanbul, Turkey, [accessed on: 03.04.2015], www.cpl.uh.edu/publication_files/C17.pdf.
- Pavlidis I., Levine J. (2002a), *Thermal Facial Screening for Deception Detection*, Proceedings of the Second Joint EMBS/BMES Conference, Houston, TX, USA, 23–26.12.2002.
- Pavlidis I., Levine J. (2002b), *Thermal Image Analysis for Polygraph Testing*, IEEE Engineering in Medicine and Biology, Vol. 21, No. 6, pp. 56–64.

- Pavlidis I., Levine J.A., Baukol P. (2000), *Thermal imaging for anxiety detection*, Proceedings 2000 IEEE Workshop on Computer Vision Beyond the Visible Spectrum: Methods and Applications, Hilton Head Island, SC, 16.06.2000, pp. 104–109.
- Polakowski H., Kastek M., Pilski J. (2011), *Analysis of facial skin temperature changes in acquaintance Comparison Question Test*, European Polygraph, Vol. 5, No. 3–4 (17–18), pp. 107–121.
- Pollina D.A., Ryan A.H. (2004), *The relationship between facial skin surface temperature reactivity and traditional polygraph measures used in the psychophysiological detection of deception: a preliminary investigation*, Polygraph, Vol. 33, No. 3.
- Pollina D.A., Dollins A.B., Senter S.M., Brown T.E., Pavlidis I., Levine J.A., Ryan A.H. (2006), *Facial Skin Surface Temperature Changes During a “Concealed Information” Test*, Annals of Biomedical Engineering, Vol. 34, No. 7, pp. 1182–1189.
- Rajoub B.A., Zwiggelarr R. (2014), *Thermal Facial Analysis for Deception Detection*, IEEE Transactions on Information Forensics and Security, Vol. 9, No. 6, p. 1015–1023.
- Salazar Lopez E., Dominguez E., de la Fuente J.M., Meins A., Iborra O., Rodriguez-Artacho M.A., Gomez Milan E. (2012), *The mental nose: thermography and the Pinocchio effect*, University of Granada, Spain.
- Shastri D., Merla A., Tsiamyrtzis P., Pavlidis I. (2009), *Imaging Facial Signs of Neurophysiological Responses*, IEEE Transactions on Biomedical Engineering, Vol. 56, No. 2.
- Shastri D., Papadakis M., Tsiamyrtzis P., Bass B., Pavlidis I. (2012), *IEEE Transactions on Affective Computing*, Vol. 3, No. 3.
- Shastri D., Tsiamyrtzis P., Pavlidis I. (2008), *Periorbital Thermal Signal Extraction and Applications*, 30th Annual International IEEE EMBS Conference, Vancouver, British Columbia, Canada, 22–24.08.2008.
- Staszal R., Wojtarowicz M., Zajac P. (2013), *Praktyczne możliwości zastosowania kamery termowizyjnej*, Studia Prawnicze. Rozprawy i Materiały, 2 (13), pp. 109–120.
- Trovillio P.V. (1938/1939), *A history of lie-detection*, The Journal of Criminal Law and Criminology, Vol. 29, No 6, p. 849.
- Tsiamyrtzis P., Dowdall J., Shastri D., Pavlidis I., Frank M.G., Ekman P. (2007), *Imaging facial physiology for the detection of deceit*, International Journal of Computer Vision, Vol. 71, pp. 197–214.
- Zhu Z., Tsiamyrtzis P., Pavlidis I. (2007), *Forehead Thermal Signature Extraction in Lie Detection*, Proceedings of the 29th Annual International Conference of the IEEE EMBS Cité Internationale, Lyon, France, 23–26.08.2007.

Jan Widacki, Michał Widacki, Jacek Antos

Preparation to Experimental Testing of the Potential from Using Facial Temperature Changes Registered with an Infrared Camera in Lie Detection*

**Подготовка до экспериментальных исследований возможности использования
изменений температуры лица, записанных с помощью тепловизионной камеры
для детекции лжи**

Key words: infrared camera in lie-detection, thermovision in detection of deception

The numerous organic changes related to emotions include fluctuations in blood pressure that can be observed and registered in alterations of pulse rate and blood pressure, and also as changes of blood flow in the bare parts of the body (face, back of the neck) that are visible with the naked eye. As changes in blood supply resulting from emotions are also accompanied by changes in the chemical composition of blood that result from the activity of endocrine glands issuing hormones into the bloodstream, changes in blood supply are also accompanied by changes of face temperature. The body temperature in a healthy human ranges from 36°C to 37°C, and its natural fluctuations may lie in the range of around 0.6°C (Konturek 2014: 495). Lowering of the body temperature is accompanied by dilation of blood vessels, sweating, and slowing or stopping of thermogenic mechanisms. Increase of temperature, on the contrary, is accompanied by the narrowing of blood vessels and increased thermogenesis (Konturek 2014: 498). Psychological and psychophysiological literature has long listed changes of body temperature as an indicator of emotions (see e.g.: Ax 1953; Hilgard 1972). Measurement of the changing temperature of individual parts of the face with classical methods, i.e. with the use of contact thermometers, was very difficult technically and impractical to the highest degree, hence it has essentially not been applied in practice. Such a measurement could have become relatively easier if a method of imaging heat radiation issued by physical bodies, i.e. thermography (thermal vision), were resorted to. This method is generally applied in an various sciences and also for practical purposes. A device used for the observation and measurement of facial temperature changes in this method is an infrared (thermal vision) camera. It can be used to detect the temperature of an observed object and its changes.

It has long been postulated that an infrared camera could be used for observation and recording of temperature changes in the human face to detect emotions, and hence also for lie detection purposes. In Poland, such a generally formulated concept was announced as early as in 1979 by Hubert KołECKI (KołECKI 1979). The use of thermal vision for lie detection seems the more attractive as the method theoretically allows to perform such a detection without physical contact, and without attaching any

* J. Widacki, M. Widacki, J. Antos, *Preparation to Experimental Testing of the Potential from Using Facial Temperature Changes Registered with an Infrared Camera in Lie Detection*, "European Polygraph" 2016, Vol. 10, No. 1 (35), p. 17–23.

sensors on the subject's body, and therefore, theoretically, also without his or her consent and even knowledge.

[This obviously gives rise to the development of new legal regulations, as all extant ones only refer to classical polygraph examinations, in which all the physiological correlates of emotions are observed with sensors installed on the body of the subject, and therefore require the subject's information and additionally consent. There is no room to discuss these questions here, as the presentation is limited to the discussion of the technical aspects of the exercise only.]

In the several recent years many experimental works aiming at obtaining this goal have been conducted, yet only a handful of general works on the subject have been published in Poland (Staszek, Wojtarowicz, Zająć 2013). The first exercises conducted were not as much experimental research projects but simple demonstration of an infrared camera and the distribution of temperatures on the face obtained with it (Polakowski, Kasteck, Pilski 2011). Available foreign literature is far richer and has recently received a broad discussion (Gołaszewski, Zająć, Widack 2015). The results of experimental studies on the use of infrared camera for lie detection conducted so far seem fairly encouraging. It seems that changes of facial temperature could be at least another physiological correlate of emotions facilitating lie detection, besides such indicators as changes in the operation of the circulatory system, the breathing patterns, and the galvanic skin response (GSR). If facial temperature changes prove to be more diagnostic than the aforementioned physiological correlates of emotions registered by a classical polygraph, they could be considered an independent measure in lie detection or used in conjunction with other contactless (remote) methods (e.g. voice change analysis, eyeball tracking, etc.).

Yet the use of an infrared camera for continuous observation and registration of facial temperature changes in parallel with the application of a polygraph that would make it possible to compare the diagnostic values of facial temperature changes to the diagnostic value of a polygraph examination requires a parallel simultaneous registration of physiological variables with a polygraph and registration of facial temperature changes with an infrared camera. Running appropriate recordings simultaneously, however, encounters a number of difficulties of technical nature. The first is that the camera shows the image of the face, in which individual colours correspond to temperatures (see: **Fig. 1**). Temperature changes are displayed as shifting areas of colour. Therefore, the first technical problem is to select a number of points on the face, monitoring temperature changes in the selected points continuously, and representing temperature changes in time in the selected points in a graphic form. In other words, the first problem was the transformation of the changing image into a graphic form.



Fig. 1. Image of the face by infrared camera.

The following one, which needs a solution, is to have a fixed infrared camera continuously observe the same previously selected points on the face of the subject, especially as the subject may, often subconsciously, jerk the head when movements that result from answering the test questions occur. These

cannot be eliminated, even if it were possible to stabilise the subject's head mechanically (e.g. with an orthopaedic device or some kind of brace), which in itself would be difficult, potentially hardly efficient, and furthermore highly impractical. The solution of the problem required writing a new piece of software that would make it possible to capture and record data from an infrared camera, and especially to present the temperature of the observed points in a graphic form along the time axis, simultaneously make it possible to denote on the graph the moment when the subject is exposed to a stimulus (test question). Moreover, the software must be constructed so as to allow continuous tracing of temperature changes in the previously selected points on the subject's face that are most diagnostic according to literature (see: Pollina *et al.* 2004; Pollina *et al.* 2006; Jain *et al.* 2012; Rajoub, Zwiggelaar 2014), and to do so, despite the movements of the subject's head and facial muscles. Application code was written in Matlab (2014 b) environment, which supports the FLIR (A655sc) infrared camera standard. Software development required also the use of Image Acquisition Toolbox support package. The data is transmitted from the camera to the computer over an Ethernet connection, the software grabs the "frames" (i.e. locations from which temperature measurement samples are taken) from both a digital and an infrared camera. The image from the digital camera is used for detection of the areas in which the infrared camera is to perform the measurements. Whenever an area has been correctly detected, the temperature is read from the corresponding "frame" of the infrared camera image. For temperature reading to be precise, the scanning area of the two cameras is shifted by the value that results from the physical distance between the two cameras mounted parallel to each other on a stand in an identical distance (120 cm) from the subject.

The processing capacity of the computer used (Dell xps 7021) proved a major limitation in the operation of the software. Initially, the number of acquisitions from the camera prohibited observation and calculation periods exceeding 15 seconds. This resulted from buffer overflow problems. To solve the problem, image transmission from the network camera to the laptop was limited to 9 kB MTU (maximum transmission unit). This provided a sufficient number of frames (12 frames per second, FPS) to conduct the necessary calculations, and extended the period of observation to 12 minutes, which is sufficient to run a polygraph test. After that time, the connection to the camera is refreshed, which clears the buffer. A following problem that called for solution was to devise an algorithm that could continuously trace selected points on the subject's face. The application makes use of Viola-Jones algorithm used for detection of objects on video streams. In this way, the data obtained is fed into a calculation sheet, which in turn generates a graph showing temperatures from the sampling points.

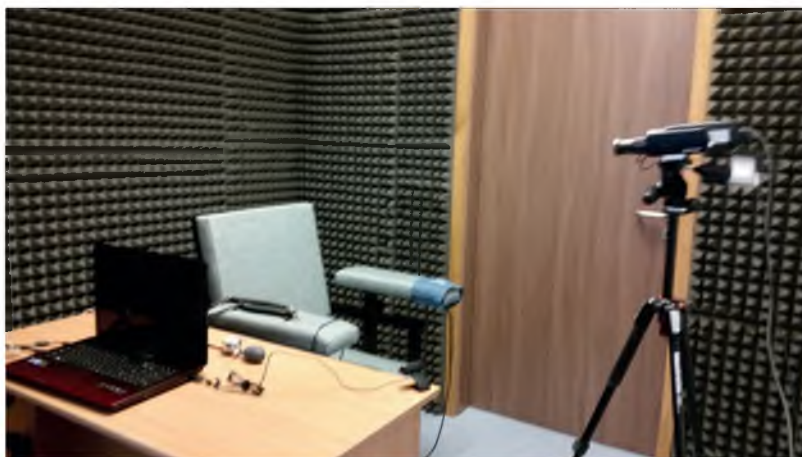


Fig. 2. Armchair, polygraph, cameras on stand, subject in the armchair.

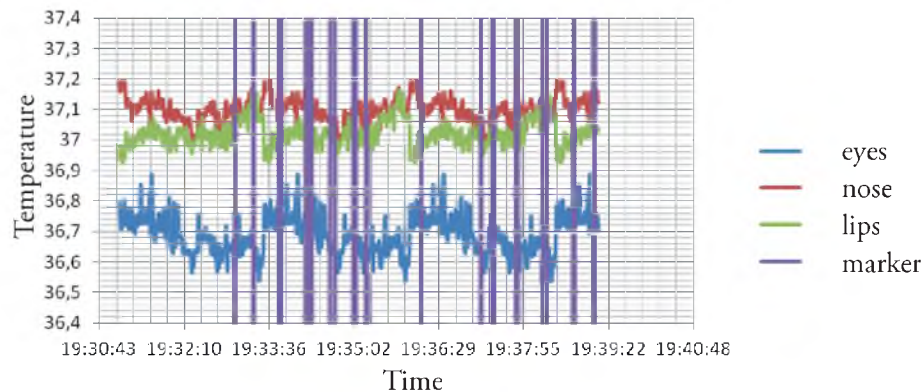


Fig. 3. A graph with the time and temperature axes with three temperature curves (eyes, nose, and lips).

In the study, we used Lafayette 4000 polygraph and a Flir Gige Vision A-600 camera. They were connected to a set composed of a Samsung R780 laptop, Dell Xps L702 X laptop, and a digital camera HIKVISION model DS.-2CD6026FHWD-A.



Fig. 4. Lafayette polygraph.



Fig. 5. Infrared camera.

Source: Images were taken from the websites of their respective manufacturers.

To allow parallel polygraph examination and tracing of facial temperature changes, an infrared camera was mounted on a stand, and focused on the face of the subject sitting in the polygraph examination room. The recording of the changes of temperature on the face was observed on a monitor by another expert in a separate room (i.e. not by the polygrapher performing the examination) and recorded. In this way, no people besides the subject and the polygrapher were present in the examination room during the polygraph examination, as required by good polygraph practice. The polygrapher and the expert observing facial temperature changes registered with an infrared camera can communicate throughout the experiment, using lan messenger 0.7 beta 4, which allows text messages exchanges between computers.

Only the overcoming of the aforementioned technical problems will make it possible to carry out the experiment aimed at comparing the diagnostic value of facial temperature changes with the diagnostic values of both complete polygraph examination and indications of its individual channels.

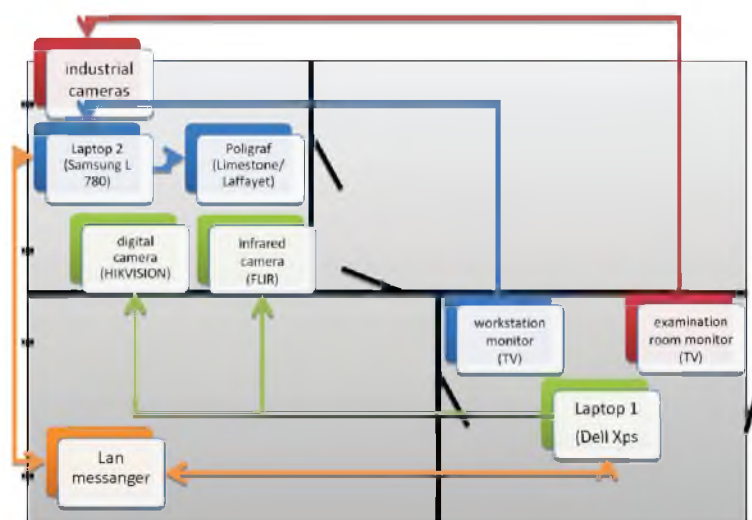


Fig. 6. Rooms where the experiment will be conducted plus listing of the basic equipment.

Source: The authors' own materials.

References

- Ax A.F. (1953), The Physiological Differentiation between Fear and Anger in Humans, *Psychosom. Med.* 1953, 15, 433–442.
- Konturek S.J. (2014) (ed.), *Fizjologia człowieka*, 2nd ed., Elsevier Urban & Partner, Wrocław 2014.
- Gołaszewski M., Zając P., Widacki J. (2015), Thermal Vision as a Method of Detection of Deception. A Review of Experience, *European Polygraph* 2015, 9, 1 (31), 5–24.
- Hilgard E.R. (1972), *Wprowadzenie do psychologii*, Warszawa 1972, 243–245.
- Jain U., Tan B., Li Q. (2012), Concealed Knowledge Identification Using Facial Thermal Imaging, Li Creative Technologies, Florham Park, NJ, USA.
- Kołecki H. (1979), *Kryminalistyczne zastosowania termografii*, Warszawa 1979.
- Polakowski H., Kastek M., Pilski J. (2011), Analysis of Facial Skin Temperature Changes in Acquaintance Comparison Question Test, *European Polygraph* 2011, 5, 3–4, 107–121.
- Pollina D.A., Ryan A.H. (2004), The Relationship between Facial Skin Surface Temperature Reactivity and Traditional Polygraph Measures used in Psychophysiological Detection of Deception: A Preliminary Investigation, *Polygraph*, 2004, 33, 3.
- Pollina D.A., Dollins A.B., Senter S.M., Brown T.E., Pavildis I., Levine J.A. Ryan A.H. (2006), Facial Skin Surface Temperature Changes During a “Concealed Information Test”, *Annals of Biomedical Engineering* 2006, 34, 7, 1182–1189.
- Rajoub B.A., Zwiggelarr R. (2014), Thermal Facial Analysis for Deception Detection, *IEEE Transactions on Information Forensics and Security* 2014, 9, 6, 1015–1023.
- Staszal R., Wojtarowicz M., Zając P. (2013), Praktyczne możliwości zastosowania kamery termowizyjnej, *Studia Prawnicze. Rozprawy i Materiały* 2013, 2 (13), 109–119.

The article is a part of Polish National Centre's (NCN), project no. DEC-2013/11/B/HS5/03856

Jan Widacki

Część VI

Aspekty prawne i etyczne

1. Wprowadzenie

Cała infrastruktura prawna dotycząca instrumentalnej detekcji kłamstwa, która rutynowo wykonywana jest wyłącznie w formie badań poligraficznych, zakłada, że badania te wykonywane są za wiedzą i zgodą badanego. Dotyczy to zarówno badań wykonywanych w sprawach kryminalnych, w ramach procesu karnego, jak i badań kadrowych wykonywanych w służbach policyjnych i specjalnych. W ostatnim okresie w Polsce przepisy prawa zostały istotnie uszczegółowione przez orzeczenie Sądu Najwyższego z 2015 roku (postanowienie Sądu Najwyższego z 29.01.2015, sygn. I KZP 25/14). Ta infrastruktura prawna została szczegółowo omówiona w całym szeregu artykułów¹. Możliwość wykonania takiego badania „bezkontaktowo”, a więc bez wiedzy i zgody badanego, wymagałoby nowych uregulowań prawnych.

2. Badania poligraficzne metodą bezkontaktową.

Problemy techniczne i prawne

Pierwsze próby instrumentalnej detekcji kłamstwa opartej na zdobyczach psychologii i fizjologii miały miejsce w końcu XIX wieku². W latach 20. XX wieku John A. Larson skonstruował dwukanałowy aparat, który rejestrował równocześnie przebieg czynności oddychania, zmiany częstotliwości tętna oraz ciśnienia krwi. Aparat ten zwany początkowo „cardio-pneumo-psychografem” (a slangowo: *Sphyggy*) z czasem zaczęto nazywać „poligrafem” lub „wykrywaczem kłamstwa” (*lie-detector*)³. Aparat został następnie udoskonalony przez Leonarda Keelera, który nie tylko dokonał technicznych, opatentowanych przez siebie udoskonaleń pneumografu i sfigmografu⁴, ale dodatkowo uzupełnił aparat psychogalwanometrem⁵.

¹ J. Widacki, *Glosa do postanowienia SN z dnia 29 stycznia 2015, I KZP 25/14*, „Państwo i Prawo” 2015, nr 8, s. 129–135; J. Widacki, A. Szuba-Boroń, *Badania poligraficzne w procesie karnym w świetle postanowienia Sądu Najwyższego z dnia 29 stycznia 2015 r., sygn. I KZP 25/14*, „Prokuratura i Prawo” 2016, 2, s. 5–16; J. Widacki, *Opinia z badań poligraficznych w procesie karnym*, „Palestra” 2017, 7–8, s. 5–10.

² Por. J. Widacki, *Historia badań poligraficznych*, Oficyna Wydawnicza AFM, Kraków 2017, s. 68 i nast.

³ *Ibidem*, s. 84 i nast. Por. też J. Widacki, *The first polygraph?*, „European Polygraph” 2016, 10, 3, s. 111–116; J. Widacki, *Poligraf czy wariograf*, „Problemy Kryminalistyki” 2016, 294, s. 35–39, 82–92.

⁴ *Ibidem*, s. 91 i nast.

⁵ Por. J. Reid, F. Inbau, *Truth and Deception. The Polygraph (Lie-detector) Technique*, Williams&Wilkins Comp., Baltimore 1977, s. 3; por. też: J. Widacki, *Poligraf czy...*, *op. cit.*, s. 91–92; a także: J. Slowik, J.P. Buckley, L. Kroeker, Ph. Ash, *Abdominal and Thoracic Respiration Recording in the Detection of Deception*, „Polygraph” 1973, 2,1, s. 12.

Przez kilkadziesiąt lat, wszystkie produkowane w świecie poligrafy rejestrowały trzy podstawowe funkcje organizmu, będące fizjologicznymi korelatami emocji: przebieg czynności oddychania, przebieg pracy układu krążenia (częstotliwość tętna i wahania ciśnienia krwi), reakcje skórno-galwaniczną. Dodatkowo niektóre modele uzupełniane były o urządzenia mierzące inne jeszcze funkcje organizmu, np. napięcie mięśni lub reakcję pletyzmograficzną. Te same funkcje rejestrowały też wprowadzane od początku lat 90. XX wieku poligrafy komputerowe. Pierwszy poligraf komputerowy skonstruowała firma Stoelting już w 1991 roku⁶, a w dwa lata później, w roku 1993 firma Lafayette wprowadziła już swoje komputerowe poligrafy do seryjnej produkcji.

Wszystkie te poligrafy, zarówno analogowe, jak i późniejsze komputerowe, wymagały założenia na ciało osoby badanej odpowiednich czujników: mankietu sfigmografu, rur pletyzmografu oraz elektrod psychogalwanometru.



Ryc. 1. Czujniki tradycyjnego poligrafu rozmieszczone na ciele badanego: dwie rury pneumografu (na klatce piersiowej i przeponie), mankiety kardiografu (zakładany na lewe ramię) oraz elektrody psychogalwanometru (zakładane na dwa niesąsiadujące palce prawej ręki)

Tak więc sama konstrukcja aparatu, wymagająca założenia na ciało badanego czujników, powodowała, że badanie mogło być wykonane wyłącznie za wiedzą osoby badanej. Nawiasem mówiąc, także sama technika badania (omówienie pytań testowych, odpowiedzi na te pytania, nieruchome siedzenie w czasie badania itp.) wymagała świadomego współdziałania badanego w czasie badania. A zatem już tylko ze względów technicznych badanie poligraficzne nie mogło być wykonane bez wiedzy badanego i jego zgody. Teoretycznie jednak, ta ostatnia mogła być w różny sposób wymuszona lub wyłudzona, co stanowi odrębny problem.

W konsekwencji badania poligraficzne mogą być wykonywane tylko za wiedzą i zgodą osoby badanej. Ten wymóg praktyczny uzyskał dodatkowo sankcję prawną. Przepisy prawa wymagają, aby badany wyraził zgodę na badanie. W Polsce kodeks postępowania karnego zezwalający na wykonanie badania poligraficznego w procesie karnym, tak w celach eliminacyjnych (art. 192 a), jak też w celach dowodowych (art. 199a) wymaga *expressis verbis*

⁶ J. Widacki, *Historia badań poligraficznych...*, op. cit., s. 136.

uzyskania zgody osoby badanej na poddanie jej takim badaniom. Przepis art. 192a § 2 kpk stanowi: „W wypadkach, o których mowa w §1, **za zgodą osoby badanej** biegły może również zastosować środki techniczne mające na celu kontrolę nieświadomych reakcji organizmu tej osoby”. Z kolei przepis art. 199a stanowi: „Stosowanie w czasie badania przez biegłego środków technicznych mających na celu kontrolę nieświadomych reakcji organizmu badanej osoby **możliwe jest wyłącznie za jej zgodą (...)**”. Przypomnieć należy, że kodeks postępowania karnego badanie poligraficzne określa opisowo jako „stosowanie środków technicznych (...) mających na celu kontrolę nieświadomych reakcji jej [tj. osoby badanej] organizmu w związku z przesłuchaniem” (art. 171 §5 pkt 2 kpk)⁷.

Poza procesem karnym badania poligraficzne są w Polsce stosowane na szeroką skalę jako element procedury rekrutacyjnej do cywilnych i wojskowych służb policyjnych i specjalnych. Podstawą prawną takiego stosowania tych badań są ustawy o tych służbach i wydane na ich podstawie akty wykonawcze. Akty te używają nazwy: „**badania poligraficzne**”⁸, najczęściej jednak stosują synonimiczne określenia „**badania psychofizjologiczne**”⁹ albo „**badania wariograficzne**”¹⁰. Wszystkie badania poligraficzne wykonywane na podstawie tych aktów prawnych wymagają także zgody osoby badanej. Natomiast przy wykonywaniu badań poligraficznych pracowników firm prywatnych, w polskim systemie prawnym dopuszczalnych¹¹ i faktycznie wykonywanych, nie ma wyraźnego prawnego obowiązku uzyskania formalnej zgody badanego. Obowiązek uprzedniego uzyskania takiej zgody albo wywodzi się z całego systemu prawa pracy, albo stosuje się go na zasadzie analogii do badań wykonywanych w sprawach karnych lub personalnych w służbach policyjnych lub specjalnych. Problemem w takich sprawach może być raczej ważność takiego oświadczenia woli, złożonego w sytuacji niekiedy dla pracownika przymusowej.

Z uprawnienia do niewyrażenia zgody na badanie poligraficzne korzysta zwykle niewielki procent osób, którym poddanie się takim badaniom zaproponowano. W praktyce amerykańskiej na przełomie lat 70 i 80 XX wieku zgody na badanie nie wyrażało od 3 do 5% takich osób. W tym czasie w Polsce odsetek odmów zgody na badanie był jeszcze mniejszy. W praktyce Zakładu Kryminalistyki Uniwersytetu Śląskiego o okresie od jesieni 1977 do końca marca 1980 roku, na 412 osób, którym zaproponowano badanie poligraficzne, zgody nie wyraziło tylko 5 osób, co stanowiło 1,2% tej grupy¹². Brak aktualnych danych

⁷ Kpk – komentarz.

⁸ Tak w ustawie z dnia 9.06.2006 o Centralnym Biurze Antykorupcyjnym, Dz.U. Nr 104, poz. 708.

⁹ Ustawa z 24.04.2003 o Agencji Bezpieczeństwa Wewnętrznego i Agencji Wywiadu, rozporządzenie Prezesa Rady Ministrów z 24.04.2003 w sprawie wzoru kwestionariusza osobowego oraz szczegółowych zasad i trybu przeprowadzania postępowania kwalifikacyjnego wobec kandydatów do służby w Agencji Wywiadu, rozporządzenie Prezesa Rady Ministrów z 29.11.2002 w sprawie wzoru kwestionariusza osobowego oraz szczegółowych zasad i trybu postępowania kwalifikacyjnego wobec kandydatów do służby w Agencji Bezpieczeństwa Wewnętrznego; rozporządzenia Ministra Obrony Narodowej z 26.07.2006 w sprawie postępowania kwalifikacyjnego wobec żołnierzy ubiegających się o wyznaczenie na stanowiska służbowe w Służbie Kontrwywiadu Wojskowego, oraz z 26.07.2006 w sprawie postępowania kwalifikacyjnego wobec kandydatów do służby w Służbie Kontrwywiadu Wojskowego; ustawa z 12.10.1990 o Straży Granicznej, ustawa z 24.08.2001 o Żandarmerii Wojskowej i wojskowych organach porządkowych; ustawa z 6.04.1990 o Policji.

¹⁰ Tak w rozporządzeniu Ministra Spraw Wewnętrznych i Administracji z 20.03.2007 w sprawie trybu i warunków ustalania zdolności fizycznej i psychicznej policjantów do służby na określonych stanowiskach lub w określonych komórkach organizacyjnych Policji; a także w rozporządzeniu Ministra Finansów z 20.12.2010 w sprawie przeprowadzenia testu sprawności fizycznej, badania psychologicznego i psychofizjologicznego funkcjonariuszy celnych.

¹¹ Por. A. Cempura, M. Widacki, *Prawna dopuszczalność pracowniczych badań poligraficznych w Polsce*, „Palestra” 2012, nr 11–12, s. 39–47.

¹² J. Widacki, *Analiza przesłanek diagnostycznych w badaniach poligraficznych*, Wydawnictwo Uniwersytetu Śląskiego, Katowice 1982, s. 87.

w tym zakresie, niemniej jak wynika z rozmów z poligraferami-praktykami, w ostatnich latach odmowy wyrażenia zgody na badania, zwłaszcza w grupie badań pracowniczych, są o wiele częstsze niż było to na przełomie lat 70. i 80. XX wieku. Może to być wynikiem zarówno większej dostępnej wiedzy na temat badań poligraficznych, jak i większej świadomości praw przysługujących jednostce.

Podsumowując ten fragment rozważań, należy stwierdzić, że obecna technika badania poligraficznego nie pozwala na przeprowadzenie takiego badania bez zgody, a więc także bez wiedzy osoby badanej. Niezależnie od tego, obowiązujące przepisy prawa, wychodząc z przesłanek etycznych i stojąc na gruncie poszanowania praw jednostek, wymagają, by badanie poligraficzne przeprowadzone być mogło jedynie za zgodą osoby badanej.

Instrumentalna detekcja kłamstwa opiera się, jak wiadomo, na założeniu, że kłamstwu, z którego wykryciem łączą się dla kłamiącego jakieś istotne konsekwencje, towarzyszą zmiany poziomu aktywności organizmu (zmiany emocjonalne oraz pewien dodatkowy wysiłek poznawczy)¹³. Czyli detekcja kłamstwa polega na obserwacji zmian poziomu aktywności organizmu w szczególnej sytuacji, jaką stwarza badanie (słuchanie pytań testowych, udzielanie na nie odpowiedzi). W czasie badania obserwuje się więc, rejestruje i ocenia fizjologiczne korelaty emocji. Obserwacja i rejestracja części z nich wymaga założenia na ciało badanego rozmaitych czujników. Część może być obserwowana i rejestrowana zdalnie, bez kontaktu urządzenia z ciałem badanego. Jak wiemy, te fizjologiczne korelaty emocji, które wykorzystywane są w badaniu poligraficznym (zmiany w pracy układu oddechowego, zmiany w pracy układu krążenia, zmiany przewodnictwa elektrycznego skóry) wymagają zainstalowania na ciele badanego odpowiednich czujników. Są jednak i takie fizjologiczne korelaty emocji, które można obserwować i rejestrować zdalnie, bez bezpośredniego kontaktu z ciałem badanego.

Do takich należą między innymi zmiany brzmienia głosu a także, od czasu gdy można je obserwować i rejestrować kamerą termowizyjną, zmiany temperatury ciała.

Potoczne doświadczenie przekonuje, że emocje odzwierciedlają się w głosie. Po głosie zdarza się nam rozpoznać nastrój mówiącego. Odróżniamy, zwłaszcza u osoby znajomej, smutny głos, głos wesoły, gniewny, niepewny, nieśmiały itp. Te zmiany głosu pod wpływem emocji mają z zasady charakter wtórny. Są efektem zmiany napięcia mięśni mechanizmu głosowego, zmiany temperatury wnętrza jamy ustnej, jej wilgotności, zmiany w przebiegu czynności oddychania itd. Istnieją techniczne możliwości mierzenia niektórych parametrów głosu, takich jak jego wysokość, natężenie czy barwa. W roku 1941 badacze amerykańscy Fay i Middleton przy pomocy „sędziów kompetentnych”, na podstawie ich subiektywnych ocen próbowali rozpoznać po głosie, czy mówiący kłamie, czy mówi prawdę. Trafność tego rozpoznania była niewiele większa od losowej¹⁴. Na początku lat 60. ubiegłego wieku Alpert i współautorzy¹⁵, posługując się aparaturą pomiarową, ustalili, że zmiany emocjonalne łatwiejsze są do wychwycenia w paśmie niskich częstotliwości (100–250 Hz), znacznie trudniejsze w pełnym paśmie częstotliwości (100–6000 Hz).

Po pokonaniu kolejnych trudności technicznych, pod koniec lat 60. udało się skonstruować urządzenie do analizy głosu pod kątem jego zmian emocjonalnych. Amerykańska firma Dektor-Counterintelligence and Security Inc. ze Springfield (Wirginia) rozpoczęła pro-

¹³ Por. *Kryminalistyka*, red. J. Widacki, wyd. 3, C.H. Beck, Warszawa 2016, s. 418 i nast.

¹⁴ P.J. Fay, W.C. Middleton, *The ability to judge truth-telling or lying from the voice as transmitted over a public address system*, „Journal of General Psychology” 1941, Vol. 24, s. 211–215.

¹⁵ H. Alpert, R.L. Kurtzberg, A.J. Friedhoff, *Transient voice changes associated with emotional stimuli*, „Archives of General Psychiatry” 1963, Vol. 8, No. 4, s. 362–365.

dukcję i sprzedaż aparatów o nazwie „Psychological Stress Evaluator”. Poszczególne modele oznaczane były PSE-1, PSE-2, PSE-101. Aparaty PSE przedstawiały w formie graficznego wykresu związane ze stresem zmiany głosu ludzkiego. W materiałach reklamowych producent wyjaśniał, że urządzenie rejestruje zmiany w głosie będące wynikiem niewielkich oscylacji mięśni aparatu mowy, wywołanych stresem. Zmiany te wedle producenta były sterowane przez centralny układ nerwowy (CUN), podczas gdy tradycyjne poligrafy rejestrują zmiany fizjologiczne, za które odpowiada autonomiczny układ nerwowy (AUN)¹⁶. Nieco później, inna amerykańska firma Law Enforcement Association Inc. z Balleville (New Jersey) rozpoczęła produkcję podobnego aparatu o handlowej nazwie Mark II – Voice Strss Analyzer (VSA). Wprowadzeniu urządzenia na rynek towarzyszyła rozbudowana akcja reklamowa. Nowe urządzenia reklamowano jako alternatywę dla poligrafów. Dla potencjalnych użytkowników szczególnie interesujące było to, że stosowanie PSE lub VSA nie wymagało instalowania na ciele badanego żadnych czujników, a więc teoretycznie pozwalało na przeprowadzenie badania nawet bez wiedzy, a tym samym zgody osoby badanej. Nic więc dziwnego, że tą techniką detekcji kłamstwa zainteresowały się Siły Zbrojne Stanów Zjednoczonych. Na ich zlecenie Joseph Kubis z Fordham University w Nowym Jorku w warunkach eksperymentu porównał wartość diagnostyczną badania wykonanego przy pomocy klasycznego 3-kanalowego poligrafu i wartość diagnostyczną badania wykonanego przy użyciu PSE i VSA. O ile w badaniach poligraficznych uzyskał on 76% trafnych wskazań, to w badaniach przy pomocy analizatorów głosu nie udało mu się uzyskać wyników nawet na poziomie statystycznej szansy. Konkluzją raportu Kubisa było stwierdzenie, że urządzenia PSE i VSA nie nadają się do rzetelnej detekcji kłamstwa i są dla wojska nieprzydatne¹⁷.

Wyniki Kubisa potwierdziły późniejsze badania Barlanda¹⁸. Na przełomie lat 90. i 2000 analizatory głosu były przedmiotem licznych badań eksperymentalnych, których wyniki publikowano w organie American Polygraph Association – „Polygraph”. Spośród tych prac na wyróżnienie zasługują w szczególności prace Barlanda, Horvatha, Haddada i współautorów¹⁹. Wykazują one, że wartość diagnostyczna badania analizatorami głosu nie tylko jest znacząco mniejsza od wartości diagnostycznej badania poligraficznego, ale nawet od wartości diagnostycznej samych wskazań psychogalwanometru. Do podobnych wniosków doszli w zasadzie autorzy polscy, którzy eksperymentowali z analizatorami głosu²⁰. Zdaje się nie ulegać wątpliwości, że sama analiza zmian głosu klasycznego badania poligraficznego nie zastąpi.

¹⁶ Por. J. Widacki, *Historia badań poligraficznych...*, op. cit., s. 137.

¹⁷ J. Kubis, *Comparison of voice analysis and polygraph as lie-detection procedure*, Fordham University, New York 1973 (prepared for US Army Land Welfare Laboratory, Contract DAAD 05-72-C 0217) (niepublikowane), cyt. za: J. Widacki, *Historia badań poligraficznych...*, op. cit., s. 138.

¹⁸ G. Barland, *Use of voice changes in the detection of deception*, paper presented on 68. Conference of the Acoustical Society of America, Los Angeles, Oct. 31, 1973 (cyt. za: J. Widacki, *Historia badań poligraficznych...*, op. cit., s. 138), a także: G. Barland, *Detection of deception in criminal suspects. A field validation study*, Department of Psychology, University of Utah 1975 (niepublikowane) cyt. za J. Widacki, *Historia badań poligraficznych...*, op. cit., s. 138.

¹⁹ G. Barland, *Use of voice changes in the detection of Deception*, „Polygraph” 2002, Vol. 31, No. 2; F. Horvath, *Experimental comparison of the psychological stress evaluator and the galvanic skin response in detection of deception*, „Polygraph” 2002, Vol. 31, No. 2; S. Wather, R. Ratley, M. Smith, *Investigation and evaluation of the voice stress technology. Final Report*; National Criminal Reference Service NCJ, No 193832 (niepublikowane).

²⁰ Por. np. J. Pietruszka, *Wykorzystanie analizatora głosu (poligrafu-wariografu głosowego) w postępowaniu karnym i stosunkach pracy* (niepublikowana praca doktorska, Uniwersytet Warszawski), 2008; por. także: M. Leśniak, B. Leśniak, M. Gramatyka, *Trafność wiarygodności wypowiedzi na podstawie analizatora głosu LVA 6.5*, [w:] *Psychologiczne i interdyscyplinarne problemy w opiniodawstwie sądowym w sprawach cywilnych*, red. J. Stanik, Katowice 2011.

Teoretycznie można rozważać, czy taką funkcję może spełnić analiza głosu połączona z innymi metodami zdalnego obserwowania i rejestrowania zmian emocjonalnych. Co najmniej dwie inne takie metody są dość zaawansowane. Pierwszą z nich jest obserwacja zmian temperatury twarzy, dokonywana zdalnie (bezkontaktowo) za pomocą kamery termowizyjnej, drugą – obserwacja ruchów gałek ocznych i odruchu źrenicznego.

Zmiany temperatury twarzy jako wskaźnik zmian emocjonalnych są przedmiotem licznych badań eksperymentalnych, także polskich²¹. Przegląd wyników badań zagranicznych z tego zakresu był opisany niedawno w literaturze polskiej²². Wyniki są zachęcające, ale metoda nie wyszła jeszcze z fazy badań eksperymentalnych i do zastosowania jej w praktyce droga jeszcze wydaje się daleka, wymagająca rozwiązania całego szeregu nawet drobnych problemów technicznych²³. Dla przykładu można wymienić choćby dwa pierwsze z brzegu. W czasie badania badany w zasadzie powinien w ogóle nie ruszać głową, a nawet ruchy mięśni twarzy towarzyszące odpowiedziom na pytania testu mogą poważnie utrudnić obserwację i rejestrację wybranego obszaru twarzy, na którym dochodzi do zmian temperatury. Również fakt, że kamera musi być umieszczona relatywnie blisko twarzy osoby badanej, czyni „bezkontaktowość” badania w dużej mierze iluzoryczną.

Wykorzystanie kamery termowizyjnej do obserwacji i rejestracji zmian temperatury twarzy jako wskaźnika zmian emocjonalnych przy detekcji kłamstwa był przedmiotem naszych badań eksperymentalnych (por. niżej Część VI).

Jeszcze mniej zaawansowana technicznie jest metoda detekcji kłamstwa poprzez zdalną obserwację ruchów gałek ocznych badanego oraz odruchu źrenicznego. Do obserwacji i rejestracji tych zmian skonstruowano specjalny aparat zwany okulografem. „Okulografia” (ang. *eye tracking*), czyli śledzenie ruchu gałek ocznych, jest techniką mającą zastosowanie do badań podstawowych w medycynie, psychologii eksperymentalnej, a w naukach stosowanych – w ergonomii czy marketingu. Jednym z kierunków wykorzystania okulografii jest detekcja kłamstwa. Podjęte w tym kierunku próby wydają się interesujące. W badaniu eksperymentalnym przeprowadzonym na Uniwersytecie w Tel Avivie, stosując technikę badania Lykken’a (Guilty Knowledge Technique) i obserwując odruch źreniczny, udało się poprawnie wskazać 50% „winnych” i 100% „niewinnych”. Dla porównania, obserwując reakcję elektrodermalną, poprawnie wskazano 55% „winnych” i 93% „niewinnych”²⁴. Podobnie zachęcające wyniki osiągnęli także inni badacze²⁵. Zdaje się nie ulegać wątpliwości, że ruchy gałek ocznych i odruch źreniczny są dobrym wskaźnikiem zmian emocjonalnych. Wykorzystanie ich obserwacji do detekcji kłamstwa jest jednak technicznie trudne, a do zdalnej dyskretnej detekcji – bardzo trudne. Współcześnie znane są trzy podstawowe typy okulografów i związane z nimi trzy podstawowe metody badania ruchu gałek ocznych i odruchu źrenicznego. Tylko jedna z nich pozwala na obserwację bezkontaktową. Polega ona na tym, że światło podczerwone odbija się od oka i następnie jest rejestrowane przez kamerę. Wykorzystuje się tu najczęściej

²¹ Por. H. Polakowski, M. Kastek, J. Pilski, *Analysis of facial skin temperature changes in acquaintance CQT*, „European Polygraph” 2011, Vol. 5, No. 3–4 (17–18), s. 107–121; J. Widacki, M. Widacki, J. Antos, *Preparation to experimental testing of the potential from using facial temperature changes registered with an infrared camera in lie-detection*, „European Polygraph” 2016, Vol. 10, No. 1 (35), s. 17–27.

²² M. Gołaszewski, P. Zajac, J. Widacki, *Thermal vision as a method of detection of deception. A review of experiences*, „European Polygraph” 2015, Vol. 9, No. 1 (31), s. 5–24.

²³ Por. J. Widacki, M. Widacki, J. Antos, *op. cit.*

²⁴ R.E. Lubow, O. Fein, *Pupillary size In response to a Visual Guilty Knowledge Test: New technique for detection of deception*, „Journal of Experimental Psychology: Applied” 1996, Vol. 2, No. 2, s. 164–177.

²⁵ Por. np. D.P. Dionisio, E. Granholm, W.A. Hillix, W.F. Perrine, *Differentiation of deception using pupillary responses as an index of cognitive processing*, „Psychophysiology” 2001, Vol. 38, No. 2, s. 205–211.

odbicie rogówki oraz środek źrenicy²⁶. Podobnie jak przy pomiarze zmian temperatury twarzy, odległość kamery od oka musi być relatywnie niewielka, co również tu czyni „bezkontaktowość” badania dość problematyczną.

Podsumowując, stwierdzić należy, że obecnie nie ma jeszcze technicznych możliwości zastąpienia poligrafu, obserwującego i rejestrującego zmiany organizmu badanego (przebieg czynności oddychania, przebieg pracy układu krążenia, zmiany w reakcji skórno-galwanicznej) metodami kontaktowymi, a więc za jego wiedzą, urządzeniem (urządzeniami), które w sposób bezkontaktowy, a więc teoretycznie bez wiedzy badanego obserwowałyby trzy inne funkcje jego organizmu.

Inaczej mówiąc, bezkontaktowa dyskretna detekcja kłamstwa w oparciu o metody psychofizjologiczne nie jest na obecnym etapie rozwoju techniki możliwa.

Zainteresowanie takimi metodami służb specjalnych wydaje się oczywiste, a badania eksperymentalne nad nimi są coraz bardziej zaawansowane. Postępowi tych badań towarzyszyć powinna refleksja prawna i etyczna, jakie granice są tu nieprzekraczalne.

Zdaje się nie ulegać wątpliwości, że w sprawach karnych i pracowniczych warunkiem dopuszczalności badania psychofizjologicznego (poligraficznego) powinna pozostać zgoda osoby badanej. A zatem badania „dyskretnie”, prowadzone bez wiedzy i zgody badanego powinny być zakazane. Czy i jakie granice prawne powinny być określone takim badaniom wykonywanym w ramach czynności operacyjno-rozpoznawczych służb specjalnych w celu zapobieżenia aktowi terroru czy przy sprawdzeniu wiarygodności ważnego źródła, pozostaje kwestią otwartą.

Wykonano w ramach projektu NCN: DEC-2013/11/B/HS 5/03856

²⁶ <https://pl.wikipedia.org/wiki/Okulografia> [dostęp: 15.10. 2017].

ANEKS E

Jan Widacki

Tezy postanowienia Sądu Najwyższego z dnia 29 stycznia 2015 r., I KZP 25/14*

Teza pierwsza:

„Uznać (...) trzeba, że niedopuszczalne jest użycie wariografu w trakcie czynności przesłuchania. Użyte przez ustawodawcę w art. §5 ust. 2 kpk wyrażenie „w związku z przesłuchaniem” rozumieć należy w ten sposób, że omawiany zakaz dotyczy nie tylko samej czynności procesowej przesłuchania, ale również czynności pozostających w bezpośredniej relacji z przesłuchaniem. Omawiany zakaz obejmuje zatem przesłuchanie z udziałem eksperta bądź samodzielne wykorzystanie wariografu przez organ przesłuchujący. Zakaz ten dotyczy także badania przeprowadzonego przez biegłego, które to badanie nie może być substytutem przesłuchania (...).

Za niedopuszczalne uznać trzeba przeprowadzenie badania bezpośrednio przed lub bezpośrednio po przesłuchaniu, kiedy mogłoby ono wpływać na swobodę wypowiedzi osoby przesłuchiwanej, stanowić rodzaj nacisku lub zagrożenia z jej punktu widzenia (...) Znaczenie wyrażenia „związek z przesłuchaniem” art. 171 §5 ust. 2 kpk nie powinno być przy tym wykładane rozszerzająco”.

Teza druga:

„Art. 192 a §2 kpk dotyczy wstępnej fazy postępowania przygotowawczego, etapu, w którym postępowanie nie toczy się jeszcze przeciwko określonej osobie – in personam, ale pozostaje postępowaniem w sprawie – in rem i konieczne jest prowadzenie czynności zmierzających do ustalenia potencjalnego podejrzanego. (...)

Art. 199a kpk stanowi podstawę przeprowadzenia badań przy użyciu wariografu w innej fazie postępowania i wobec innych osób, aniżeli wskazane w art. 192a kpk. Skoro w zdaniu drugim art. 199a kpk wyłączono stosowanie art. 199 kpk, a ten odnosi się wprost do oskarżonego (a tym samym podejrzanego – art. 71 §3 kpk), to uznać trzeba, że przepis art. 199a kpk dotyczy postępowania karnego w fazie, w której toczy się ono przeciwko konkretnej osobie, a więc po przedstawieniu zarzutów, ze stadium postępowania sądowego włącznie. Przepis ten stanowić może podstawę przeprowadzenia dowodu z opinii biegłego, który stosuje środki techniczne mające na celu kontrole nieświadomych reakcji organizmu osoby badanej, a więc w czasie badania stosuje wariograf (poligraf)”.

Teza trzecia:

„Dowód z opinii biegłego jest dowodem o charakterze pomocniczym, pośrednim (używa się także określeń specyficznym lub uzupełniającym) i nie może zastępować samodzielnych dowodów (...). Jeśli opinia biegłego odpowiada warunkom określonym w art. 200 kpk, podlegać będzie swobodnej ocenie sądu, tak jak każdy inny dowód (...).”

* J. Widacki, Glosa do postanowienia SN z dnia 29 stycznia 2015 r., I KZP 25/14, „Państwo i Prawo” 2015, nr 8, s. 129–135.

Glosa do postanowienia Sądu Najwyższego z dnia 29 stycznia 2015 (I KZP 25/14)

1. Głosowane postanowienie całościowo reguluje stosowanie poligrafu (wariografu) w polskim procesie karnym. Oprócz wymienionych wyżej tez, które będą przedmiotem rozważań glosy, zawiera ono też szereg dalszych wskazówek dotyczących sposobu przeprowadzania dowodu z badań poligraficznych (badanie wykonane wyłącznie za zgodą badanego, uprzedzonego dodatkowo o konsekwencjach wynikających z art. 199a kpk, wykonane możliwie w najwcześniejszej fazie postępowania). Mamy tu zatem do czynienia po raz pierwszy z tak szeroko przedstawionym stanowiskiem Sądu Najwyższego w sprawie badań poligraficznych.

2. W pierwszej głosowanej tezie, Sąd Najwyższy wyjaśnił, jak jego zdaniem należy rozumieć ograniczenie zastosowania badania poligraficznego wynikające z zawartego w art. 171 § 5 ust. 2 kpk wyrażenia „w związku z przesłuchaniem”. Co zatem znaczy, że badanie poligraficzne, dopuszczalne w polskiej procedurze karnej po nowelizacji kpk w 2003 roku (Dz. U. 2003, Nr 17, poz. 155) dodającej art. 192 a i 199 a, nie może być wykonane „w związku z przesłuchaniem”? Zdaniem Sądu Najwyższego, z brzmienia tego fragmentu przepisu wynika, że po pierwsze, badanie nie może być wykonane przez przesłuchującego w trakcie wykonywania czynności procesowej jaką jest przesłuchanie, ale musi być wykonane przez biegłego w formie ekspertyzy. Wydaje się to oczywiste. Co więcej, wydawało się oczywiste jeszcze przed nowelizacją kpk w 2003 roku (por. wyrok Sądu Apelacyjnego w Krakowie z dnia 19. sierpnia 1999 w sprawie II AKa 147/99 z glosą aprobowaną: J. Widackiego, *Palestra* 2000, 2–3, 251). W głosowanym orzeczeniu Sąd Najwyższy idzie jednak dalej. Zdaniem Sądu, z wyrażenia „w związku z przesłuchaniem” wynikają dalej jeszcze idące ograniczenia. Jednym z nich jest zakaz przesłuchania z udziałem eksperta z zakresu badań poligraficznych. Stanowisko to jest zasadniczo słuszne, nie ulega bowiem wątpliwości, że nie wolno przesłuchiwać z udziałem eksperta z zakresu badań poligraficznych. Ale jak się wydaje, zakazu tego nie trzeba wyprowadzać z art. 171 § 5 ust. 2 kpk. Dla takiego przesłuchania brak po prostu w polskiej procedurze podstaw prawnych. Przesłuchanie z udziałem biegłego (ale tylko lekarza lub psychologa) jest przewidziane w art. 192 §2 kpk, jednak wyłącznie w sytuacji, gdy istnieje wątpliwość co do stanu psychicznego świadka, jego rozwoju umysłowego, zdolności postrzegania lub odtwarzania spostrzeżeń. Przesłuchanie z udziałem biegłego przewiduje też art. 185 a kpk ale dotyczy to jedynie przesłuchania pokrzywdzonego małoletniego, przy czym przepis wyraźnie wskazuje, że ma to być biegły psycholog. Krótko mówiąc, niezależnie od treści art. 171 §5 ust. 2 kpk, przesłuchanie z udziałem biegłego z zakresu badań poligraficznych nie jest możliwe, z braku podstaw prawnych dla takiej czynności. Sąd Najwyższy generalnie słusznie zwraca uwagę, że badanie poligraficzne wykonane przez biegłego nie może być substytutem przesłuchania. Trzeba jednak pamiętać, że konieczną częścią składową badania poligraficznego jest wywiad przed testowy (rozmowa przed testowa). W czasie tego wywiadu niezbędne jest między innymi ustalenie, co osoba badana (podejrzany, świadek, osoba faktycznie podejrzewana lub znajdująca się w kręgu osób faktycznie podejrzewanych) wie (lub twierdzi, że wie) na temat przestępstwa w sprawie którego ma być badana, konieczne jest też omówienie z osobą badaną pytań, które będą jej zadane w czasie badania. Co więcej, ustawodawca w zdaniu drugim art. 199a kpk wyłączył w stosunku do osoby badanej poligraficznie działanie przepisu art. 199 kpk. A zatem oświadczenia podejrzanego (oskarżonego) dotyczące zarzucanego mu czynu, złożone wobec biegłego z zakresu badań poligraficznych, inaczej niż takie same oświadczenia złożone wobec innych biegłych (np. lekarzy, psychologów) mogą stanowić dowód. Już to niebezpiecznie zbliża badanie poligraficzne do procesowej czynności przesłuchania, czyniąc w tym zakresie jej surogat. Zatem stanowisko Sądu Najwyższego w tej kwestii należy odczytywać tak, że badanie poligraficzne wykonywane przez biegłego nie może być bardziej substytucyjne wobec przesłuchania, niż wynika to z samej istoty tego badania. Tak więc na przykład, biegłemu nie wolno namawiać badanego do przyznania się, nie może wypytywać o szczegóły sprawy nie mające znaczenia dla badania poligraficznego, a być może przydatne w śledztwie, wypytywać o inne osoby i t.p. Sąd Najwyższy słusznie w głosowanym orzeczeniu zauważa, że „znaczenie wyrażenia „związek z przesłuchaniem” nie powinno być wykładane rozszerzająco” Twierdzi też, że niedopuszczalne jest

przeprowadzenie badania „bezpośrednio przed lub bezpośrednio po przesłuchaniu”. Dodaje jednak: „kiedy mogło by ono wpłynąć na swobodę wypowiedzi osoby przesłuchiwanej, stanowić rodzaj nacisku lub zagrożenia z jej punktu widzenia”. Tę wypowiedź można w zasadzie interpretować dwojako. Pierwsza możliwa interpretacja jest taka, że badanie poligraficzne nie może być wykonane w bezpośrednim sąsiedztwie czasowym przesłuchania, ale tylko, gdy w konkretnej sytuacji (do której odnosi się słowo „kiedy”) mogło by to stanowić formę nacisku na osobę przesłuchiwaną. Druga możliwa teoretycznie interpretacja, prowadzić mogłaby do stwierdzenia, że bezpośrednie sąsiedztwo czasowe tych dwóch czynności (przesłuchania i badania) zawsze stanowi formę nacisku (lub co najmniej grozi takim naciskiem) na przesłuchiwanego (słowo „kiedy” rozumiane by było jako „zawsze” „kiedy to”) i takie czasowe sąsiedztwo jest z zasady niedopuszczalne. Jednak z uwagi na wspomnianą dyrektywę nierozszerzającej wykładni raczej nie wchodzi w grę. Niedopuszczalne zatem wydaje się stworzenie takiego bloku czynności podobnego do tego, który realizowany jest przy okazji okazania (przesłuchanie, okazanie, przesłuchanie). Zatem niedopuszczalna jest stworzenie bloku: przesłuchanie przez organ procesowy, badanie poligraficzne wykonane przez biegłego, przesłuchanie przez organ procesowy na okoliczność badania.

Tak więc stanowisko Sądu Najwyższego w kwestii rozumienia zawartego w art. 171§5 ust. 2 kpk wyrażenia „w związku z przesłuchaniem” można podsumować następująco:

- 1) Badanie poligraficzne w procesie musi być wykonane przez biegłego w formie ekspertyzy; z uznania tego zdanie wynika uznanie kilku następnych zdań, jako jego logicznych konsekwencji. A więc biegły taki musi być powołany przez organ procesowy w formie postanowienia o powołaniu biegłego, musi mieć wiadomości specjalne w zakresie badań poligraficznych, z badań ma być sporządzona opinia spełniająca wymogi art. 200 kpk, a nadto badanie nie może być wykonane przez organ procesowy w ramach przesłuchania, nie można łączyć roli biegłego i organu procesowego, czyli ta sama osoba nie może raz badać jako biegły, innym razem jako organ procesowy przesłuchiwać osobę badaną.
- 2) Badanie nie może być wykonane przez biegłego w trakcie przesłuchania prowadzonego przez organ procesowy.
- 3) Niedopuszczalne jest tworzenie bloku czynności: przesłuchanie, badanie poligraficzne odnoszące się do przesłuchania, przesłuchanie odnoszące się do badania poligraficznego.
- 4) Unikać należy sytuacji zacierających granice między badaniem a przesłuchaniem (np. pobieranie zgody na badanie w czasie przesłuchania, nawiązywanie w badaniu poligraficznym do odbytego przesłuchania, dopytywanie w czasie badania o kwestie będące przedmiotem przesłuchania i faktyczne „dosłuchiwanie” badanego).

3. W drugiej głosowanej tezie Sąd Najwyższy zajął stanowisko w sporze toczącym się w orzecznictwie i doktrynie, co do tego, czy badanie poligraficzne jest dopuszczalne (oczywiście za zgodą badanego) tylko w fazie postępowania *in rem* i tylko w celu określonym w art. 192 a kpk („w celu ograniczenia kręgu osób podejrzanych”) czy też jest ono dopuszczalne także w stosunku do podejrzanego (oskarżonego) czy świadka.

To pierwsze stanowisko, ograniczające dopuszczalność badań poligraficznych w procesie jedynie w ograniczenia kręgu podejrzanych i zezwalające na poddanie takim badaniom, oczywiście za ich zgodą, jedynie osób o nieokreślonym jeszcze statusie procesowym, w fazie postępowania *ad rem*, a zakazujące badań podejrzanego (oskarżonego) lub świadka było wyraźnie reprezentowane w doktrynie (por. np. Stefański R., [w:] Bartoszewski J., Gardocki L., Gostyński Z., Przyjemski S.M., Stefański R., Zabłocki S., Kodeks postępowania karnego – Komentarz, t. I, Warszawa 2004 s. 880, Grzegorzczak T., procesowe aspekty badań poligraficznych w świetle znowelizowanych przepisów procedury karnej, Palestra, 2003, 11–12, 146). Stanowisko to występowało też w orzecznictwie sądów (por. postanowienie Sądu Najwyższego z 9.02. 2010 II KK 198/09, OSN KW 2010, 5, 47 oraz Biuletyn Sądu Najwyższego 2010, 4, 18; wyrok Sądu Apelacyjnego w Katowicach z 25.04.2007, II Aka 67/07 Legalis). Z satysfakcją należy przyjąć, że w głosowanym wyroku Sąd Najwyższy potwierdził stanowisko Sądu Najwyższego

wyrażone w postanowieniu z 17.10. 2012, IV KK 237/12 (niepubl.), że art. 199 a kpk, dopuszcza możliwość posłużenia się badaniem poligraficznym wykonanym przez biegłego w ramach ekspertyzy „za zgodą oskarżonego”. Jeśli badany może być oskarżony, to tym samym takie badanie po pierwsze odbyć się musi już po fazie *in rem*, gdy badany *ex definitione* ma już określony status procesowy, a celem badania, w tym etapie postępowania, w szczególności gdy już jest oskarżony (podejrzany) nie może być w sposób oczywisty eliminacja osób podejrzewanych. Za dopuszczalnością badań poligraficznych nie tylko w celach eliminacyjnych, ale także dowodowych, już po fazie postępowania *in rem*, Sąd Najwyższy i Sądy Apelacyjne wypowiadały się już w przeszłości kilkakrotnie (por. wyroki Sądu Najwyższego z 20.01. 2010, IV KK 364/09 Legalis, Sądu Apelacyjnego w Krakowie z 26.04. 2005 II Aka 264/04 KZS 2005,5, 43; wyrok Sądu Apelacyjnego w Lublinie z 25.03.2010, II Aka 15/10, Legalis).

Takie stanowisko prezentowane było także w doktrynie (por. np. L. Paprzycki: Przedmiot opinowania psychiatrycznego i psychologicznego w postępowaniu karnym – wybrane zagadnienia [w:] Aktualne problemy prawa i procesu karnego. Księga ofiarowana prof. Janowi Grajewskiemu, red. M. Płachta, Gdańsk 2003, Widacki J. (red.): Badania poligraficzne w Polsce, Oficyna Wydawnicza AFM, Kraków 2014, s. 87). Co więcej, analiza spraw karnych za lata 2003–2012, w których wykorzystano badania poligraficzne pokazuje, że zdecydowana większość przeprowadzonych w procesach karnych badań poligraficznych dotyczyła już podejrzanych (a nie osób bez określonego jeszcze statusu procesowego) a celem badania nie była eliminacja osób faktycznie podejrzanych, ale dostarczenie dowodu w postępowaniu. Badania poligraficzne w celach eliminacyjnych, o jakich mowa w art. 192 a kpk, stanowią w praktyce polskiej znikomy odsetek wszystkich badań poligraficznych wykonywanych w sprawach karnych (por. Widacki J., (red.), Badania poligraficzne w Polsce, Oficyna Wydawnicza AFM, Kraków 2014, s.88).

4. W trzeciej głosowanej tezie, Sąd Najwyższy powtórzył powszechny w orzecnictwie i literaturze pogląd, że dowód z opinii biegłego z badań poligraficznych jest „dowodem pomocniczym, pośrednim (...) i nie może zastępować samodzielnych dowodów”. A nadto, przypomniał, że „jeśli opinia biegłego odpowiada warunkom określonym w art. 200 kpk, podlegać będzie swobodnej ocenie sądu, tak jak każdy inny dowód”. To ostatnie zdanie zawiera prawdy oczywiste i jako takie zdaje się nie budzić żadnych wątpliwości. Nie można tego powiedzieć o zdaniu poprzednim. Już w 1976 roku Sąd Najwyższy uznał, że „badania poligraficzne mają jedynie charakter pomocniczy i nie mogą stanowić samodzielnego dowodu, dającego podstaw do konkretnych ustaleń” (wyrok Sądu Najwyższego z 25.09. 1976, II KR 171/76, niepublikowany, choć wielokrotnie cytowany w literaturze, z glosą W. Daszkiewicza i M. Jeż-Ludwichowskiej, Państwo i Prawo 1979, 5;). Zwracano już wówczas uwagę, że polska procedura karna nie zna takiej kategorii dowodu jak „dowód pomocniczy”, a o ważności i roli danego dowodu w każdym konkretnym postępowaniu decyduje sąd na zasadzie swobodnej oceny dowodów (por. np. S. Waltoś: Badania poligraficzne w Polsce w świetle przepisów prawa, poglądów Sądu Najwyższego i nauki [w:] materiały konferencji Wykorzystanie badań poligraficznych w sprawach karnych, Katowice 1978 W. Daszkiewicz, M. Jeż-Ludwichowska: Badania wariograficzne w polskim procesie karnym, Nowe Prawo 1979, 3; J. Widacki: Analiza przesłanek diagnozowania w badaniach poligraficznych, Wyd. UŚ, Katowice 1982, s.8). Przyjąć można, że nazwanie dowodu „pomocniczym” służy jedynie wskazaniu, że ocena tego dowodu powinna być szczególnie ostrożna, a oprócz tego dowodu, potrzebne są jeszcze inne. Skądinąd wiadomo, że ocena dowodów dokonywana przez sąd faktycznie jest oceną dokonywaną między innymi poprzez pryzmat innych dowodów zebranych w sprawie i w ich kontekście. Zatem, żaden dowód, w tym dowód z opinii z badań poligraficznych, nie może stanowić samodzielnej podstawy do wydania wyroku.

Sąd Najwyższy w głosowanym orzeczeniu po pierwsze „dowód pomocniczy” uznaje za synonim „dowodu pośredniego”, po drugie uznaje, że taki dowód (pomocniczy, pośredni) nie może zastępować samodzielnych dowodów. Zatem „dowód pomocniczy” jest przeciwieństwem „dowodu samodzielnego”. Wypada tu tylko powtórzyć, że polska procedura nie zna podziału dowodów na „samodzielne” i „pomocnicze”. Natomiast doktryna, zna podział na dowody „pośrednie” i „bezpośrednie” (por. M. Cieślak : zagadnienia dowodowe w procesie karnym, t. I Wyd. Prawnicze, Warsza-

wa 1955, s. 82 i nast.; A. Gaberle: Dowody w sądowym procesie karnym, Oficyna Wolters Kluwer business, Kraków 2007, s. 44 i nast.) Ten ostatni podział znany jest nauce procesu karnego także w innych krajach. W procesie anglosaskim na przykład, dowód bezpośredni to „direct evidence”, dowód pośredni, to „circumstantial evidence” (por. np. Ingram J.L.,: Criminal evidence, Elsevir, Amsterdam, Boston, Heidelberg, London, New York, Oxford, Paris, San Diego, San Francisco, Singapore, Sydney, Tokyo 2012).

Dowody bezpośrednie, to takie dowody, które odnoszą się wprost do faktu głównego w procesie. Na przykład przyznanie się do zabójstwa przez podejrzanego, zeznanie świadka, który widział zabójcę w trakcie czynu to dowody bezpośrednie. Dowody pośrednie, to z kolei takie, które pozwalają na przyjęcie ustaleń, z których dopiero można wnioskować coś na temat faktu głównego. Na przykład ustalenie, że na odzieży podejrzanego znajduje się ślad biologiczny pochodzący od ofiary. Otóż badanie poligraficzne dostarcza dowodu pośredniego.

Prawidłowo sformułowana konkluzja opinii z badań poligraficznych wykonanych techniką pytań kontrolnych powinna zawierać zdanie: **„badany na pytania testu reagował tak, jak zwykle reagują osoby, które na te pytania odpowiadają nieściszerze, t.j. albo świadomie kłamią, albo ukrywają fakt posiadania informacji, o które są w badaniu pytani”**. Słowo „zwykle” należy rozumieć w kontekście wartości diagnostycznej badania poligraficznego, a ta w różnych badaniach jest szacowana na 80–95%. Czyli „zwykle”, rozumieć trzeba, że na 100 kłamiących, tak jak badany zareagowało by od 80 do 95. Albo innymi słowy: na 100 prawdomównych, tak jak badany zareagowało by od 5 do 20 (por. J. Widacki: Results of polygraph examination. Direct or circumstantial evidence? European Polygraph, 2014, 8, 2 (28), 61–67). To, czy konkretny badany należy do reagującej typowo większości czy do reagującej nietypowo mniejszości, rozstrzygnąć musi organ procesowy, oceniając wynik badania poligraficznego w kontekście innych dowodów. Ocena dowodu z badań poligraficznych dokonywana przez sąd jest identyczna, jak w przypadku wszystkich dowodów pośrednich. Mamy tu do czynienia, z logicznego punktu widzenia z rozumowaniem (wnioskowaniem) redukcyjnym, czyli wnioskowaniem ze znanego następstwa (reakcja na pytania krytyczne) o niepewnej (bo tylko jednej z możliwych) przyczynie (racji). Zatem badanie poligraficzne, jak każde badanie wykonane przez biegłego w ramach ekspertyzy, dostarcza dowodu pośredniego.

Jako dowód pośredni badanie poligraficzne nie nadaje się do samodzielnej weryfikacji wyjaśnień oskarżonego, będących ze swej natury dowodem bezpośrednim. Inaczej, samo stawało by się nie tylko dowodem bezpośrednim, ale takim dowodem bezpośrednim, który ma a priori określoną wartość, niepodlegającą z jednej strony ocenie sądu, z drugiej zaś strony wyręczającym sąd w ocenie innego dowodu.

Na marginesie warto zauważyć, że w większości polskich podręczników kryminalistyki, badania poligraficzne (wariograficzne) umieszczane są właśnie, jakby na przekór temu stanowisku w rozdziałach poświęconych przesłuchaniom i prezentowane są jako metody weryfikacji zeznań i wyjaśnień.

Podsumowując. Opinia biegłego z badań poligraficznych dostarcza dowodu pośredniego. Dowód pośredni nie może służyć do samodzielnej weryfikacji dowodu bezpośredniego, tak samo, jak nie może służyć do samodzielnego ustalenia faktu głównego w procesie. Dowód z opinii z badań poligraficznych nie jest bardziej „pośredni” (ani bardziej „pomocniczy”) niż jakikolwiek inny dowód z opinii biegłego.

Jan Widacki, Anna Szuba-Boroń

Badania poligraficzne w procesie karnym w świetle postanowienia Sądu Najwyższego z dnia 29 stycznia 2015 (I KZP 25/14)*

Wprowadzenie

Proces karny nie jest jedynym miejscem wykorzystania badań poligraficznych, znacznie więcej przeprowadza się ich współcześnie w procedurach przedzatrudnieniowych i kontrolnych realizowanych w służbach policyjnych i specjalnych, a także w ramach czynności operacyjno-rozpoznawczych wykonywanych przez te służby. Jednak wykorzystanie poligrafu w procesie karnym, szczególnie w śledztwie było historycznie pierwszym, a po dziś dzień pozostaje najbardziej dyskusyjnym. Przedmiotem toczącej się już od ponad 50 lat dyskusji, zapoczątkowanej w 1963 roku, gdy Paweł Horoszowski po raz pierwszy zastosował poligraf w procesie¹ była sama dopuszczalność badań poligraficznych w procesie karnym, warunki ewentualnej dopuszczalności, dopuszczalne techniki badań, wreszcie charakter uzyskanego w ten sposób dowodu. Od czasu, gdy nowelizacja kodeksu postępowania karnego z 2003 roku, wprowadzająca przepisy art. 192a i 199a² przesądziła o dopuszczalności badań w procesie karnym, dyskusja zasadniczo ograniczyła się do tego, czy badanie poligraficzne może być wykorzystane tylko w celach eliminacyjnych, jak wynika z dosłownej wykładni art. 192a kpk, czy też może być wykorzystane w celach dowodowych. Czy, jak wynika z przyjęcia pierwszego stanowiska, badaniom można poddać tylko osoby o nieokreślonym jeszcze statusie procesowym (osoby podejrzwane, faktycznie podejrzwane), czy też, jak wynika z drugiego stanowiska, można badaniom poddać także podejrzanego a nawet oskarżonego lub świadka, a w konsekwencji jaki charakter będzie miał taki dowód?

Orzecznictwo tak Sądu Najwyższego jak i sądów apelacyjnych nie było w tej kwestii konsekwentne.

Tak np. Sąd Apelacyjny w Katowicach w postanowieniu z dnia 25.04.2007 (II AKo 67/07, Legalis) stwierdził m.in.: „zwrócić przy tym należy uwagę na to, że ustawodawca dopuszcza możliwość posłużenia się poligrafem (wariografem) wyłącznie w celu określonym w art.192a §1 kpk, bowiem przepis art. 171§5 pkt 2 kpk zakazuje posługiwania się tego rodzaju środkami technicznymi w związku z przesłuchaniem, a uzyskane w ten sposób informacje nie mogą stanowić dowodu (art. 171§7 kpk)”.

* J. Widacki, A. Szuba-Boroń, *Badania poligraficzne w procesie karnym w świetle postanowienia Sądu Najwyższego z dnia 29 stycznia 2015 r., sygn. I KZP 25/14*, „Prokuratura i Prawo” 2016, t. 2, s 5–16.

¹ J. Widacki, *Historia badań poligraficznych w Polsce*, [w:] J. Widacki (red.), *Badania poligraficzne w Polsce*, Oficyna Wydawnicza AFM, Kraków 2014, s. 35.

² Ustawa o zmianie ustawy – Kodeks postępowania karnego, ustawy – przepisy wprowadzające kodeks postępowania karnego, ustawy o świadku koronnym oraz ustawy o ochronie informacji niejawnych z dnia 10 stycznia 2003 r. (Dz.U. z 2003 r., Nr 17, poz. 155).

Również Sąd Najwyższy w postanowieniu z dnia 9.02.2010r.³ stanął na stanowisku, że wyniki badań poligraficznych „służą jedynie celom eliminacyjnym i w żadnym wypadku nie mogą być dowodem przeciwko osobie badanej, gdyby uzyskała ona status podejrzanego (oskarżonego)”. Stanowisko orzecznictwa, ograniczające możliwość stosowania badań poligraficznych w procesie karnym jedynie do celów eliminacyjnych podzielane było przez część doktryny⁴.

Nie brak jednak orzeczeń, które bezpośrednio lub co najmniej pośrednio akceptują dowodowe wykorzystanie badań poligraficznych, próbując przy tym określić charakter tego dowodu („dowód niesamoistny”, „pomocniczy” co będzie niżej przedmiotem oddzielnych rozważań).

Do orzeczeń takich zaliczyć można wyroki Sądu Apelacyjnego w Krakowie z 15.01.2004 (II Aka 24/03) i z 24.04.2005 (II Aka 264/04, KZS 2005, nr 5 s. 43), wyrok Sądu Najwyższego z 28.10.2004 (III KK 51/04, OSNKW 2005, 1), wyrok Sądu Apelacyjnego w Szczecinie z 30.10.2008 (II Aka 94/08, Biul. SA w Szczecinie 2009, 1, 29-34), wyrok Sądu Apelacyjnego w Lublinie z 23.03.2010 (II Aka 15/10, LEX 583679), wyrok Sądu Najwyższego z 20.01.2010 (IV KK 364/09, LEX 560612).

Faktem jest, że mimo prawnej dopuszczalności badań poligraficznych w procesie karnym, mimo stworzenia kilku placówek policyjnych wykonujących takie badania (Centralne Laboratorium Kryminalistyczne Policji w Warszawie, KP w Białymstoku, Bydgoszczy, Łodzi i Katowicach), placówki Żandarmerii Wojskowej (Wydział Badań Psychofizjologicznych), Centralnego Biura Antykorupcyjnego, Straży Granicznej a nadto mimo możliwości wykorzystania w razie potrzeby w śledztwach ekspertów z ABW badań poligraficznych dla celów procesowych wykonuje się bardzo mało, mniej niż w 200 sprawach rocznie⁵. Biorąc pod uwagę liczbę prowadzonych każdego roku spraw karnych, łatwo wyliczyć, że jedno badanie przypada na kilka tysięcy postępowań. Świadczy to ewidentnie o niewykorzystanym potencjale tak aparaturowym jak i ludzkim. W warunkach amerykańskich, taka liczba badań jaką wykonuje się w Polsce dla potrzeb procesu karnego stanowi roczną normę dla jednego, maksimum dwóch poligraferów.

Taki stan rzeczy ma wiele przyczyn. Jedną z nich jest jak się wydaje niski poziom wiedzy o badaniach poligraficznych i ich możliwościach wśród prawników w tym także niestety wśród autorów podręczników kryminalistyki, które bardziej dezinformują niż informują na temat badań poligraficznych⁶, a także wśród prowadzących śledztwa prokuratorów. Kolejną przyczyną jest niski poziom badań realizowanych przez polskich ekspertów, których opinie z badań są z zasady nierozstrzygające, (podczas gdy w praktyce np. amerykańskiej zakłada się, że liczba badań nierozstrzygających nie jest większa niż 10%), a w konsekwencji dla śledztwa nieprzydatne. Dochodzi do tego nieudolność organizacyjna, powodująca, że zalecane przez reguły sztuki przeprowadzenie badania w możliwie jak najszybszym czasie po zatrzymaniu osoby mającej być badaną, najlepiej w czasie pierwszych 48 godzin, gdy badanie jest najbardziej efektywne, nawet gdy jest to zatrzymanie planowane, jest organizacyjnie trudne a często niewykonalne! Jednym z istotnych czynników utrudniających dotychczasową praktykę stosowania badań poligraficznych w procesie karnym było także niejasne stanowisko Sądu Najwyższego i sądów apelacyjnych co do kilku podstawowych kwestii. Należały do nich wspomniana wyżej wątpliwość, czy badanie poligraficzne można prowadzić tylko w celach eliminacyjnych, w stosunku do osób bez określonego jeszcze statusu procesowego, wyłącznie w fazie *in rem* postępowania przygotowawczego, czy także w celach dowodowych, w stosunku do podejrzanego w fazie *in personam*. A jeśli

³ Postanowienie Sądu Najwyższego z dnia 9 lutego 2010 r., II KK 198/09 [publ. OSNKW 2010, 5/47, Biul. SN 2010, 4, 18].

⁴ Por. np. R. Stefański, [w:] J. Bartoszewski, L. Gardocki, Z. Gostyński, S.M. Przyjemski, R.A. Stefański, S. Zabłocki, *Kodeks postępowania karnego, Komentarz*, t. I, Warszawa 2004, s. 880; T. Grzegorzczak, *Procesowe aspekty badań poligraficznych w świetle znowelizowanych przepisów procedury karnej*, Palestra 2003, 11–12, s. 146–151.

⁵ J. Widacki, *Polygraph examination in criminal cases. Current Polish practice. A critical study*, European Polygraph 2012, 6, 4, s. 249–256.

⁶ Por. J. Widacki, *Badania poligraficzne w polskich podręcznikach kryminalistyki*, [w:] J. Widacki (red.), *Badania poligraficzne w Polsce*, Oficyna Wydawnicza AFM, Kraków 2014, s. 235–268.

badanie dopuszczalne jest także w tym drugim przypadku, to w jakiej formie badanie ma być prowadzone, przez kogo i na koniec jaki walor dowodowy ma wynik badania.

Postanowienie Sądu Najwyższego z 29. stycznia 2015 roku (I KZP 25/14) po raz pierwszy w praktyce tego Sądu całościowo reguluje stosowanie poligrafu w polskim procesie karnym i ma szansę w znacznym stopniu uporządkować dotychczasową praktykę orzecniczą a także praktykę wykonywania badań poligraficznych w śledztwie.

Dopuszczalność badań poligraficznych w celach eliminacyjnych i dowodowych

We wskazanym wyżej postanowieniu, Sąd Najwyższy jednoznacznie wypowiedział się za dopuszczalnością badań poligraficznych zarówno w celach eliminacyjnych, dla zawężenia kręgu osób podejrzanych jak i w celach dowodowych.

W tym pierwszym przypadku, podstawą prawną przeprowadzenia takich badań jest art. 192a kpk. Zdaniem Sądu Najwyższego, *„art. 192a kpk dotyczy wstępnej fazy postępowania przygotowawczego, etapu w którym postępowanie nie toczy się jeszcze przeciwko określonej osobie – in rem, ale pozostaje postępowaniem w sprawie – in personam i konieczne jest prowadzenie czynności zmierzających do ustalenia potencjalnego podejrzanego”*.

Dla badań, realizowanych w celach dowodowych, przez poddanie im podejrzanego (ewent. świadka) Sąd Najwyższy znajduje podstawę w art. 199a kpk. Sąd Najwyższy słusznie uważa, że *„art. 199a kpk stanowi podstawę do przeprowadzenia badań przy użyciu wariografu w innej fazie postępowania i wobec innych osób aniżeli wskazane w art. 192a kpk. Skoro w zdaniu drugim art. 199a kpk wyłączono stosowanie art. 199 kpk, a ten odnosi się wprost do oskarżonego (a tym samym podejrzanego – art. 71§3 kpk) to uznać trzeba, że art. 199a kpk dotyczy postępowania karnego w fazie, w której toczy się ono przeciwko konkretnej osobie, a więc po postawieniu zarzutów, ze stadium postępowania sądowego włącznie. Przepis ten stanowić może podstawę przeprowadzenia dowodu z opinii biegłego, który stosuje środki techniczne mające na celu kontrolę nieświadomych reakcji organizmu osoby badanej, a więc w czasie badania stosuje wariograf (poligraf)”*.

Konsekwencje praktyczne wynikające z faktu, że badanie poligraficzne nie może być wykonywane „w związku z przesłuchaniem”

Z brzmienia przepisu art. 171 § 5 pkt 2 kpk wynika, że badanie poligraficzne nie może być realizowane „w związku z przesłuchaniem”. Przepis ten, sam w sobie niejasny, różnie mógł być interpretowany. Przed nowelizacją kpk w 2003 roku interpretowany był dwojako. Jedni wywodzili z niego zakaz stosowania poligrafu w procesie karnym. Tak na przykład przepis ten interpretował Sąd Apelacyjny w Białymstoku w wyroku z 14.05.1998 roku⁷, stwierdzając, że *„nowy kodeks postępowania karnego w sposób jednoznaczny a art. 171§ 4 pkt 2 i §4 [art. 171 w brzmieniu sprzed nowelizacji w 2003] rozstrzygnął kwestię stosowania hipnozy i poligrafu w postępowaniu karnym, przyjmując, iż są to metody objęte zakazem dowodowym”*.

Natomiast Sąd Najwyższy w postanowieniu z dnia 21.12. 1998⁸, w tym samym stanie prawnym zaliczył wynik badania poligraficznego do dowodów. Również Sąd Apelacyjny w Krakowie w wyroku z dnia 19.08.1999⁹ uznał, że art. 171 § 4 pkt 2 i § 4 kpk nie wprowadzają zakazu wykorzystania badania poligraficznego w celach dowodowych w procesie karnym, ale jedynie

⁷ Wyrok Sądu Apelacyjnego w Białymstoku z dnia 14 maja 1998r.; II AKa 25/98, [publ. OSA 1999, 9, 64; OSAB 1998, 2, 26].

⁸ J. Widacki, *Glosa do postanowienia Sądu Najwyższego z dnia 21 grudnia 1998 r., IV KO 101/98*, Palestra 1999, 3–4, s. 237.

⁹ J. Widacki, *Glosa do wyroku Sądu Apelacyjnego w Krakowie z dnia 19 sierpnia 1999 r., II Aka 147/99*, Palestra 2000, 2–3, s. 251.

wprowadzają zakaz wykonywania badania poligraficznego w ramach czynności przesłuchania. Natomiast przepis ten nie sprzeciwia się wykonaniu badania poligraficznego w ramach oddzielnej ekspertyzy, przez biegłego.

Po nowelizacji kodeksu postępowania karnego w 2003 roku wprowadzającej wielokrotnie już przywoływane art. 192a i 199a kpk, nie było już wątpliwości, że przepis art. 171 kpk nie zabrania stosowania poligrafu, nie wprowadza w tym względzie zakazu dowodowego, jednak wątpliwość, jak należy rozumieć, zwrot „w związku z przesłuchaniem”. Czy zwrot ten jest jednoznaczny ze zwrotem „w ramach przesłuchania”, czy znaczy coś innego. W konsekwencji, czy każde wykonanie badania poligraficznego przez biegłego w ramach zleconej mu postanowieniem ekspertyzy jest *ex-definitione* „bez związku z przesłuchaniem”? – jak przyjął w swym wyroku z 19.08.1999 Sąd Apelacyjny w Krakowie (sprawa II Aka 147/99)?

Sąd Najwyższy w omawianym postanowieniu stwierdził, że „użyte przez ustawodawcę w art. 171 §5 pkt 2 wyrażenie „w związku z przesłuchaniem” rozumieć należy w ten sposób, że omawiany zakaz dotyczy nie tylko samej czynności procesowej przesłuchania, ale również czynności pozostających w bezpośredniej relacji z przesłuchaniem”.

Sąd Najwyższy rozumie to tak: „Omawiany zakaz obejmuje zatem przesłuchanie z udziałem eksperta bądź samodzielne wykorzystanie wariografu przez organ przesłuchujący. Zakaz ten dotyczy także badania przeprowadzonego przez biegłego, które to badanie nie może być substytutem przesłuchania. Za niedopuszczalne uznać trzeba przeprowadzenie badania bezpośrednio przed lub bezpośrednio po przesłuchaniu, kiedy mogło by ono wpływać na swobodę wypowiedzi osoby przesłuchiwanej, stanowić rodzaj nacisku lub zagrożenia z jej punktu widzenia (...) Znaczenie wyrażenia „związek z przesłuchaniem” art. 171§ 5 pkt 2 kpk nie powinno być przy tym wykładane rozszerzająco”.

Stanowisko Sądu Najwyższego narzuca więc praktyce stosowania poligrafu w procesie karnym pewne rygory, rzutując tym samym na taktykę postępowania w śledztwie.

Przeanalizujmy kolejno wszystkie te rygory. To, że badanie ma wykonać biegły w ramach ekspertyzy wydaje się oczywiste. Z uznania tego zdania wynika konieczność uznania kilku kolejnych. Przede wszystkim biegłego musi powołać specjalnym postanowieniem organ procesowy (art. 193 kpk¹⁰). Badanie musi być wykonane przez biegłego z listy Sądu Okręgowego lub imiennie powołanego ad hoc (art. 195 kpk¹¹). Po wykonaniu badania biegły ma obowiązek sporządzenia opinii, która musi spełniać wymogi formalne określone przez kpk (art. 200 § 2 kpk¹²).

Sąd Najwyższy słusznie uważa, że biegły z zakresu badań poligraficznych nie może brać udziału w przesłuchaniu. Jest to, jak się wydaje wymóg oczywisty, wynikający nie tylko z art. 171§ 5 pkt 2 kpk, ale przede wszystkim z tego, że udział biegłego w przesłuchaniu jest czynnością przewidzianą w kpk, ale został ograniczony tylko do biegłego psychologa lub lekarza i tylko w ściśle określonych sytuacjach (art. 192 § 2 kpk, 185a kpk). Zatem niezależnie od treści art. 171 § 5 pkt 2 kpk, przesłuchanie z udziałem biegłego z zakresu badań poligraficznych nie jest możliwe, z powodu braku podstaw prawnych dla takiej czynności.

Sąd Najwyższy uważa także, że zakazane jest także takie prowadzenie badania poligraficznego, by mogło ono być substytutem przesłuchania. To słuszna uwaga. Badaniem poligraficznym nie można zastępować przesłuchania. Jednak konieczną częścią składową badania poligraficznego jest wywiad przedtestowy (rozmowa przedtestowa). Wywiad pełni wielorakie funkcje niezbędne dla badania poligraficznego (por. np. *Kryminalistyka* 2.wyd., red. J. Widacki, C.H. Beck, Warszawa 2012, s. 379 i n.). W trakcie wywiadu niezbędne jest jednak ustalenie, co osoba badana (podejrzany, świadek, osoba faktycznie podejrzewana lub znajdująca się w kręgu osób faktycznie podejrzewanych) wie, lub twierdzi, że wie na temat przestępstwa, w sprawie którego ma być badana. W czasie tego wywiadu omawia się też z osobą badaną pytania, które zamierza się jej

¹⁰ Ustawa z dnia 6 czerwca 1997 r. Kodeks postępowania karnego (Dz.U.1997r., Nr 89, poz. 555).

¹¹ *Ibidem*.

¹² *Ibidem*.

zadać w czasie badania, ewentualnie pod wpływem jej wypowiedzi modyfikuje się te pytania. Już to zbliża badanie poligraficzne do przesłuchania. Mało tego, ustawodawca w art. 199a zd. 2 kpk wyłączył w stosunku do osoby badanej poligraficznie działanie art. 199 kpk. Zatem jej oświadczenia, w szczególności, gdy jest to osoba podejrzana, złożone wobec biegłego z zakresu badań poligraficznych, dotyczące zarzucanego jej czynu złożone wobec biegłego z zakresu badań poligraficznych, w przeciwieństwie do takich samych oświadczeń złożonych wobec innych biegłych (np. lekarzy czy psychologów) mogą stanowić dowód. Już to, z woli ustawodawcy zbliża badanie poligraficzne wykonane w ramach ekspertyzy do procesowej czynności przesłuchania, stając się w tym zakresie jej surogatem.

Zatem, trzeba to interpretować w sposób następujący. Badanie poligraficzne wykonywane przez biegłego w ramach ekspertyzy nie może być bardziej substytucyjne wobec czynności procesowej przesłuchania, niż wynika to z samej istoty tego badania. W szczególności biegłemu nie wolno namawiać badanego do przyznania się („bo badanie i tak to wykaże”), wypytywać o szczegóły sprawy nie mające bezpośredniego znaczenia dla badania poligraficznego, a może być przydatne w śledztwie, wypytywać o inne osoby, o które badany nie będzie pytany w czasie badania. Z drugiej strony, to nie przesłuchujący powinien pobierać od przesłuchiwanego oświadczenie o wyrażeniu zgody na badanie (jak to obecnie bywa praktykowane zwłaszcza w prokuraturach wojskowych), ale biegły, przed rozpoczęciem badania.

Sąd Najwyższy jak widać poszedł jeszcze dalej. Uznał, że niedopuszczalne jest przeprowadzenie badania „bezpośrednio przed lub bezpośrednio po przesłuchaniu, kiedy mogłoby ono wpływać na swobodę wypowiedzi osoby przesłuchiwanej, stanowić rodzaj nacisku lub zagrożenia z jej punktu widzenia”. Słowo „kiedy” może tu być rozumiane dwojako. Gdyby je rozumieć jako „kiedy to”, w sensie „zawsze”, należałoby przyjąć, że zdaniem Sądu Najwyższego bliskość czasowa badania poligraficznego i przesłuchania wpływa na swobodę wypowiedzi i stanowi rodzaj nacisku lub zagrożenia z punktu widzenia osoby przesłuchiwanej. Ale słowo „kiedy” może też być tu rozumiane inaczej, jako „wtedy gdy”. Gdyby tak je rozumieć, zakaz wprowadzony przez Sąd Najwyższy nie miałby charakteru bezwzględnego. Zabronione byłoby wykonywanie badania poligraficznego bezpośrednio przed i bezpośrednio po przesłuchaniu, ale tylko wtedy, gdy w danej konkretnej sytuacji mogło by to ograniczać swobodę wypowiedzi osoby przesłuchiwanej, stanowić w jej odbiorze formę nacisku. Z uwagi na zawartą w tymże postanowieniu Sądu Najwyższego dyrektywę, że „znaczenie wyrażenia „związek z przesłuchaniem” w art. 171§5 pkt 2 kpk nie powinno być (...) wykładane rozszerzająco”, druga z przedstawionych możliwych interpretacji wydaje się właściwa.

Z tego orzeczenia Sądu Najwyższego wynikają istotne zalecenia dla praktyki stosowania poligrafu w śledztwie, regulujące postępowanie zarówno organu procesowego prowadzącego postępowanie jak i dla biegłych. Spróbujmy je podsumować.

1. Badanie poligraficzne może być zlecone zarówno w celu eliminacyjnym, dla zawężenia kręgu osób faktycznie podejrzewanych, w fazie postępowania *in rem* (podstawą zarządzenia takich badań jest wtedy art. 192a kpk), jak i w celach dowodowych, w późniejszych fazach postępowania, gdy badaniom można poddać podejrzanego (oskarżonego) lub świadka (podstawą jest wtedy art. 199a kpk).
2. Badanie poligraficzne może być przeprowadzone wyłącznie za zgodą osoby badanej, przez biegłego w ramach ekspertyzy.
3. Badanie poligraficzne, jako ekspertyza, ma być maksymalnie jak to jest tylko możliwe oddzielone od innych czynności procesowych, w szczególności od przesłuchania. Wynikają z tego wskazania szczegółowe: biegły nie może wyřęcać przesłuchującego, czyniąc z badania, a szczególnie z tej jego fazy jaką jest wywiad przedtestowy, substytutu przesłuchania; to biegły, a nie przesłuchujący ma pobierać od badanego oświadczenie w przedmiocie zgody na badanie; biegły nie powinien być obecny w czasie przesłuchania, a przesłuchujący w czasie badania poligraficznego; nie powinno się tworzyć bloku czynności: przesłuchanie, w którym zapowiedziano badanie poligraficzne – badanie poligraficzne – przesłuchanie na okoliczność

badania; nie powinno się w badaniu poligraficznym bardziej niż jest to niezbędne z uwagi na wymogi techniki badania, nawiązywanie do treści przesłuchania itp.

4. Z badania poligraficznego sporządzona ma być opinia czyniąca zadość wymogom z art. 200 kpk.

Wydaje się, że wynikające z orzeczenia Sądu Najwyższego dyrektywy postępowania przy badaniu poligraficznym zostały wyrażone jasno i powinny być respektowane zarówno przez organy procesowe zlecające takie badania, jak i przez biegłych.

Wartość dowodu z badań poligraficznych

O ile poprzednie uwagi Sądu Najwyższego adresowane są głównie do organów procesowych powołujących biegłych i zlecających im przeprowadzenie badań poligraficznych, to dokonana ocena wartości dowodu z takich badań i określenie rodzaju takiego dowodu adresowane jest przede wszystkim do sądów, ułatwiając im ocenę tego dowodu i jego właściwe wykorzystanie w procesie.

W omawianym orzeczeniu Sąd Najwyższy stwierdził, że *„dowód z opinii biegłego [wykonującego badania poligraficzne] jest dowodem o charakterze pomocniczym, pośrednim (używa się też określeń specyficznym lub uzupełniającym i nie może zastępować samodzielnych dowodów”,* dodając, że *„jeśli opinia biegłego odpowiada warunkom określonym w art. 200 kpk, podlegać będzie swobodnej ocenie sądu, tak jak każdy inny dowód”.*

Nazwanie dowodu z badań poligraficznych *„dowodem pomocniczym”,* który w konsekwencji *„nie może stanowić samodzielnego dowodu dającego podstawę do konkretnych ustaleń”* nastąpiło po raz pierwszy blisko 40 lat temu, w wyroku Sądu Najwyższego z 25.09.1976¹³.

Komentując ten wyrok Daszkiewicz i Jeż-Ludwichowska¹⁴ zwrócili uwagę, że obowiązujące prawo karnoprocesowe nie zna takich pojęć jak „dowód główny” i „dowód pomocniczy”. A zatem dowód z badań poligraficznych „jest poszlaką i musi być traktowany, jak inne dowody pośrednie”, a przez te ostatnie rozumie się dowody, które bezpośrednio nie odnoszą się do faktu głównego.

Podobnie co do charakteru dowodu z badań poligraficznych wypowiedział się Sąd Apelacyjny w Krakowie w wyroku z 15.01.2004 roku¹⁵ twierdząc, że *„dowód z badania poligraficznego nie jest samoistnym dowodem sprawstwa przestępstwa, a wskazuje jedynie na prawdopodobieństwo takiej możliwości, bo dowodzi istnienia śladu pamięciowego przeżyć związanych z popełnieniem przestępstwa”.* Wprowadzając kolejne nieznanie polskiej procedurze pojęcie „dowodu samoistnego”, objaśnia równocześnie, że rozumie go w zasadzie jako „dowód pośredni”.

Również Sąd Najwyższy w wyroku z 1.02.2008 roku¹⁶ słusznie uznał, że dowód z badań poligraficznych nie nadaje się do weryfikacji prawdziwości wyjaśnień oskarżonego, albowiem badanie prowadzi jedynie do *„rejestracji nieświadomych reakcji organizmu, – dowód taki w sposób bezsporny nie dowodzi [prawdomówności lub kłamstwa], a podlega on przy tym ocenie jak każdy inny dowód i to w kontekście także pozostałych dowodów”.*

W literaturze¹⁷ przyjmuje się, że nazwanie przez Sąd Najwyższy, na gruncie zasady swobodnej oceny dowodów, dowodu „pomocniczym” ma pewien sens praktyczny. Taki dowód wymaga szczególnie wnikliwej oceny, nie może być jedynym dowodem w sprawie, nie może samodzielnie stanowić podstawy do konkretnych ustaleń. Podobnie Sąd Najwyższy podszedł do ekspertyzy

¹³ Sprawa: II KR 171/76.

¹⁴ W. Daszkiewicz, M. Jeż-Ludwichowska, *Glosa do wyroku Sądu Najwyższego z 25.08. 1976, II KR 171/76*, Państwo i Prawo 1979, 34, 5, s. 178.

¹⁵ Wyrok Sądu Apelacyjnego w Krakowie z dnia 15 stycznia 2004 r.; II Aka 24/03 [publ. KZS 2004/5/29].

¹⁶ Wyrok Sądu Najwyższego z dnia 1 lutego 2008 r.; V KK 231/07 [publ. OSNKW 2008/4/27, Biul. PK 2008/4/13, Biul. SN 2008/3/14].

¹⁷ Por. np. *Badania poligraficzne w Polsce*, red. J. Widackiego, Oficyna Wydawnicza AFM, Kraków 2014, s. 106.

osmologicznej, stwierdzając w wyroku z 14.01.2002 r., że „dowód osmologiczny nie daje jak dotąd takiego przekonania, jakie może wyniknąć np. z badań daktyloskopijnych czy badań DNA (...) nie dyskredytując go, należy zatem dowód ten in concreto poddawać wnikliwej analizie z uwzględnieniem pozostałego materiału dowodowego”¹⁸. Jak widać w przekonaniu Sądu Najwyższego dowody, choć oceniane po myśli art. 7 kpk swobodnie, mają jednak a priori określoną pewną wartość. Niektóre dają większy stopień pewności (np. daktyloskopia, badania DNA) i niektóre takie, które takiej pewności nie dają i dlatego nie mogą samodzielnie stanowić podstaw do „konkretnych ustaleń”. Do tych ostatnich Sąd Najwyższy zaliczył m.in. ekspertyzę osmologiczną. Można więc sądzić, że podobnie traktuje badania poligraficzne i tak należy rozumieć je jako „dowód pomocniczy”. Jednak w omawianym tu postanowieniu z 29.01.2015 roku (I KZP 25/14) Sąd Najwyższy dowód z badań poligraficznych nazywa „dowodem o charakterze pomocniczym, pośrednim”, który „nie może zastępować samodzielnych dowodów”.

Znów należałoby przypomnieć, że polska procedura karna nie zna „dowodów samodzielnych”. Przeciwnieństwem dowodu „pośredniego” jest „dowód bezpośredni”. Ten ostatni podział dowodów, na bezpośrednie i pośrednie znany jest doktrynie¹⁹.

Taki podział dowodów znany też jest w procesach karnych innych krajów, rozróżniających dowody bezpośrednie (*direct evidence*) i pośrednie (*circumstantial evidence*)²⁰.

Można więc sądzić, że dla Sądu Najwyższego w omawianym orzeczeniu „dowód pomocniczy” to synonim „dowodu pośredniego”. Byłoby to zgodne z tym, jak rozumieli zwrot „dowód pomocniczy” w orzeczeniu Sądu Najwyższego z 25.09.1977 (II KR 171/76) Daszkiewicz i Jeż-Ludwichowska²¹. należy więc przypomnieć, że dowody bezpośrednie, to takie, które odnoszą się wprost do faktu głównego w procesie (np. przyznanie się podejrzanego, zeznanie naocznego świadka zdarzenia). Natomiast dowody pośrednie, to takie, które pozwalają na przyjęcie ustaleń, z których dopiero na podstawie rozumowania redukcyjnego można wnioskować coś o fakcie głównym. Na przykład wynik ekspertyzy genetyczno sądowej, stwierdzający na miejscu zdarzenia ślad DNA podejrzanego.

Takie stanowisko Sądu Najwyższego zasługuje na akceptację. Badanie poligraficzne z całą pewnością nie dostarcza dowodu bezpośredniego, ale jedynie dowodu pośredniego²².

Jeśli badanie poligraficzne dostarcza dowodu pośredniego, („badany reaguje na pytania krytyczne testu tak, jak reagują osoby, które na takie pytania odpowiadają nieszczerze, tj. kłamią lub zatajają fakt posiadania informacji”), to badanie poligraficzne nie nadaje się do samodzielnej, bezpośredniej weryfikacji wyjaśnień podejrzanego (oskarżonego), które są dowodem pośrednim. Jego wynik, na równi z innymi dowodami może stanowić jedynie kontekst w którym organ procesowy dokona swobodnej oceny wyjaśnień podejrzanego (oskarżonego). Niestety, w opozycji do tego stanowiska w większości polskich podręczników kryminalistyki badania poligraficzne (wariograficzne) umieszczane są w rozdziałach poświęconych przesłuchaniu i prezentowane są jako metody weryfikacji zeznań i wyjaśnień.

Przyjęcie stanowiska, że badanie poligraficzne dostarcza dowodu pośredniego, jest równoznaczne z przyjęciem stanowiska, że jest to dowód identyczny, jak wszystkie inne dowody pośred-

¹⁸ Wyrok Sądu Najwyższego z dnia 14 stycznia 2002 r., II KKN 465/00; [publ. LEX nr 559934].

¹⁹ Por. M. Cieślak, *Zagadnienia dowodowe w procesie karnym*, t. I, Wydawnictwo Prawnicze, Warszawa 1955, s. 82; A. Gaberle, *Dowody w sądowym procesie karnym*, Oficyna Wolters Kluwer business, Kraków 2007, s. 44; W. Daszkiewicz, M. Jeż-Ludwichowska, *Glosa do wyroku Sądu Najwyższego z 25.09.1979, II KR 171/76*, Państwo i Prawo 1979, 5, s. 178.

²⁰ Por. np. J.L. Ingram, *Criminal evidence*, Elsevier, Amsterdam, Boston, Heidelberg, London, New York, Oxford, Paris, San Diego, San Francisco, Singapore, Sydney, Tokio 2012, s. 98–104; T.F. Kiely, *Forensic evidence: science and the criminal law*, CRC Press, Boca Raton, London, New York, Washington 2001, s. 38–39, 169, 173, 277; K. Inman, N. Rudin, *Principles and practice of criminalistics. The profession of forensic science*, CRC Press, Boca Raton, London, New York, Washington 2001, s. 101–102, 110, 158.

²¹ W. Daszkiewicz, M. Jeż-Ludwichowska, *op. cit.*, s. 178.

²² Por. J. Widacki, *Results of polygraph examination: direct or circumstantial evidence*, European Polygraph 2014, Vol. 8, No. 2 (28), s. 61–67.

nie, jakimi są wyniki wszelkich badań identyfikacyjnych wykonanych w procesie. To, że wartość diagnostyczna badania poligraficznego nie jest mniejsza, niż wartość diagnostyczna innych rutynowo stosowanych w procesie metod identyfikacyjnych zostało już udowodnione przed laty²³. Co więcej, jak pokazują analizy porównawcze²⁴, podobna jest wartość diagnostyczna badań poligraficznych w relacji do badań stosowanych w medycynie, natomiast w porównaniu z badaniami psychologicznymi wykazano wyższą trafność badań poligraficznych.

Streszczenie

Badania poligraficzne są wykorzystywane w Polsce w sprawach karnych jedynie incydentalnie, podczas gdy ich wartość diagnostyczna nie różni się od innych stosowanych rutynowo w śledztwie metod identyfikacyjnych. Postanowienie Sądu Najwyższego z dnia 29 stycznia 2015 jako pierwsze całościowo reguluje zasady zastosowania poligrafu w polskim procesie karnym. W artykule przeanalizowano zarówno dopuszczalność przeprowadzania badań poligraficznych w procesie karnym, jak również wyjaśniono narzucone przez stanowisko Sądu Najwyższego rygory zastosowania poligrafu.

²³ Por. J. Widacki, *Wartość diagnostyczna badania poligraficznego i jej znaczenie kryminalistyczne*, Wyd. UJ, Kraków 1977; J. Widacki, F. Horvath, *An experimental investigation of the relative validity and utility of the polygraph technique and three others common methods of criminal identification*, Journal of Forensic Sciences 1978, 23, 3, s. 596–601.

²⁴ P.E. Crewson, *Comparative Analysis of Polygraph with Other Screening and Diagnostic Tools*, 2001, Polygraph 32, s. 57–85.

Jan Widacki

Opinia z badań poligraficznych w procesie karnym*

1. Kodeks postępowania karnego po nowelizacji w roku 2003¹ dopuszcza *expressis verbis* wykorzystanie badania poligraficznego w procesie karnym, tak w celach eliminacyjnych (art. 192a§2 kpk), jak i w celach dowodowych (art. 199a kpk)². Z przepisu art. 171§5 pkt 2 kpk, a także art. 199a kpk wynika jednoznacznie, że badanie takie musi być wykonywane przez biegłego w ramach ekspertyzy, nie może natomiast być traktowane jako część przesłuchania lub jego forma, ani prowadzone równolegle z przesłuchaniem. Co w art. 171§5 pkt 2 kpk znaczy zwrot „w związku z przesłuchaniem” wyjaśnił niedawno Sąd Najwyższy w postanowieniu z dnia 29. Stycznia 2015 roku (sygn. I KZP 25/14)³ Sąd Najwyższy stwierdził mianowicie, że „użyte przez ustawodawcę w art. 171§5 pkt 2 kpk wyrażenie „w związku z przesłuchaniem” rozumieć należy w ten sposób, że omawiany zakaz dotyczy nie tylko samej czynności procesowej przesłuchania, ale również czynności pozostających w bezpośredniej relacji z przesłuchaniem”⁴. Sąd Najwyższy wyjaśnia to następująco: „omawiany zakaz obejmuje zatem przesłuchanie z udziałem eksperta bądź samodzielne wykorzystanie wariografu przez organ przesłuchujący. Zakaz ten dotyczy także badania przeprowadzonego przez biegłego, które to badanie nie może być substytutem przesłuchania. Za niedopuszczalne uznać trzeba przeprowadzenie badania bezpośrednio przed lub bezpośrednio po przesłuchaniu, kiedy mogłoby ono wpływać na swobodę wypowiedzi osoby przesłuchiwanej, Stanowić rodzaj nacisku lub zagrożenia z jej punktu widzenia (...) Znaczenie wyrażenia „związek z przesłuchaniem” użytego w art. 171 §5 pkt 2 kpk nie powinno być przy tym wykładane rozszerzająco”⁵. Uwzględniając sama istotę badania poligraficznego, którego częścią składową jest wywiad przed testowy, czynność z natury rzeczy podobna do przesłuchania, można powiedzieć, że **badanie poligraficzne wykonywane przez biegłego w ramach ekspertyzy nie może być bardziej „substytucyjne” wobec czynności procesowej przesłuchania niż wynika to z samej istoty tego badania, w tym także wywiadu przed testowego**. W szczególności więc biegłemu nie wolno namawiać badanego do przyznania się, wypytywać o szczegóły nie mające znaczenia dla badania, ale mogące mieć znaczenie dla śledztwa, itp.⁶. Z kolei to nie przesłuchujący powinien, jak to ma czasem miejsce w praktyce, pytać przesłuchiwanego czy wyrazi zgodę na badanie poligraficzne, pobierać od niego taką zgodę na piśmie. Te czynności powinny być zarezerwowane dla biegłego⁷.

* J. Widacki, *Opinia z badań poligraficznych w procesie karnym*, „Palestra” 2017, nr 7–8.

¹ Ustawa z dnia 10 stycznia 2003 r. (Dz.U. Nr 17, poz. 155).

² Kodeks nie używa nazwy „badanie poligraficzne”, ale badania takie nazywa opisowo: „stosowaniem środków technicznych mających na celu kontrolę nieświadomych reakcji organizmu osoby badanej” (por. art. 171§5 pkt 2, art. 192 a§2 kpk, art. 199 a kpk).

³ OSNKW 2015, nr 5, poz. 38.

⁴ *Ibidem*.

⁵ *Ibidem*.

⁶ Por. J. Widacki, A. Szuba-Boroń, *Badania poligraficzne w procesie karnym w świetle postanowienia SN z dnia 29.01. 2015 (sygn. I KZP 25/14)*, Prokuratura i Prawo 2016, 2, s. 11.

⁷ *Ibidem*.

2. We wspomnianym postanowieniu Sąd Najwyższy słusznie uznał, że dowód z badań poligraficznych dostarcza „dowodu pomocniczego”. Podobnie dowód ten określał wyrok Sądu Najwyższego z 25.09.1976 (sygn. II KR 171/76, niepublikowany, wielokrotnie omawiany w literaturze). Jednak tym razem, Sąd Najwyższy wyjaśnił, że nieznanemu kodeksowi ani doktrynie „dowód pomocniczy”, jest synonimem „dowodu pośredniego”. Ten ostatni, wprawdzie nie jest też znany kodeksowi, ale jest od dawna znany doktrynie procesu karnego⁸. Jest to podział niezwykle użyteczny, znany także literaturze zagranicznej⁹. Przypomnijmy: dowody bezpośrednie, to dowody dotyczące wprost faktu głównego. Dowody pośrednie, to takie, których związek z faktem głównym wymaga przeprowadzenia wnioskowań. Jeśli na przykład oskarżony przyzna się do zabójstwa – to jego wyjaśnienia stanowią dowód bezpośredni (prawdziwy lub fałszywy). Jeśli świadek zezna, że widział jak podejrzany zabijał, to jego zeznanie (prawdziwe lub fałszywe) stanowi dowód bezpośredni. Jeśli natomiast na odzieży podejrzanego stwierdzono DNA ofiary, to uzyskujemy dowód pośredni. Trzeba dopiero wykazać, że jest mało prawdopodobne, aby ślad ten powstał w sytuacji innej niż zabójstwo. W literaturze¹⁰ podkreślano już, że wynik badania poligraficznego dostarcza dowodu pośredniego (będzie jeszcze o tym mowa niżej).

3. Skoro badanie poligraficzne może być wykonywane wyłącznie w ramach ekspertyzy przez biegłego, co uznajemy za pewnik, to wynika z tego kilka dalszych kwestii. Kwestia pierwsza, to kogo można powołać jako biegłego, uznając, że ma on wiadomości specjalne z zakresu badań poligraficznych? Sprawa bynajmniej nie jest prosta, w Polsce istnieje wiele firm oferujących badania poligraficzne. Na listach biegłych sądowych widnieją także biegli z tej dyscypliny. Nie wdając się w tym miejscu w rozważania, na jakiej podstawie prezesi poszczególnych Sądów Okręgowych dokonywali wpisu na listę biegłych z tej specjalności, wystarczy może odnotować, że większość tych wpisanych na listy biegłych nie ukończyła żadnego profesjonalnego przeszkolenia, nie składała żadnego egzaminu, nie legitymuje się żadnym certyfikatem. Wobec braku w Polsce procedur certyfikujących biegłych z tej dziedziny, jedynymi biegłymi z certyfikatem, czyli sprawdzonymi kwalifikacjami, są biegli legitymujący się certyfikatem American Polygraph Association. Jest ich w Polsce kilkunastu, w większości zatrudnionych w organach bezpieczeństwa państwa. Dwa istniejące w Polsce stowarzyszenia, a mianowicie Polskie Towarzystwo Badań Poligraficznych i Stowarzyszenie Polskich Poligraferów, udzielają rekomendacji osobom, które w ich przekonaniu dysponują wiadomościami specjalnymi z zakresu badań poligraficznych, umieszczając listy tych osób na swoich stronach internetowych. Większość działających w Polsce poligraferów, reklamujących swe usługi w Internecie, niestety także figurujących czasem na listach biegłych różnych Sądów Okręgowych nie posiada ani certyfikatu American Polygraph Association, ani rekomendacji polskich stowarzyszeń zajmujących się badaniami poligraficznymi. Są to zatem osoby o niepotwierdzonych wiadomościach specjalnych w rozumieniu art. 193 kpk, czego powołujące ich w charakterze biegłych organy procesowe powinny być świadome, podobnie jak sądy, oceniające wartość dowodową przedkładanych im wyników badań. Z faktu, że badanie poligraficzne wykonywane może być w procesie karnym tylko w formie ekspertyzy wynikają nadto następujące kwestie: o dopuszczeniu takiego dowodu (z opinii biegłego) należy wydać postanowienie o treści określonej w art. 194 kpk. W miarę potrzeby można biegłemu udostępnić akta (w zakresie niezbędnym do wydania opinii) – art. 198 kpk. Przed wykonaniem badania należy osobę badaną zapytać o zgodę na udział w badaniu (art. 192 a §2 kpk, art. 199a kpk). Po badaniu biegły zobowiązany jest złożyć opinię o określonej treści i formie (art. 200 kpk). Postanowienie o powołaniu biegłego, zgodnie z art. 194 pkt 2 kpk powinno określać „przedmiot i zakres ekspertyzy, ze sformułowaniem, w miarę potrzeby, pytań szczegółowych”. Realizacja tego punktu art. 194 kpk w dużej mierze przesądza o formie i treści

⁸ Por. np. M. Cieślak, *Zagadnienia dowodowe w procesie karnym*, t. I, Wydawnictwo Prawnicze, Warszawa 1955, s. 82; A. Gaberle, *Dowody w sądowym procesie karnym*, Wydawnictwo Wolters Kluwer, Kraków 2007, s. 44; S. Waltoś, *Proces karny. Zarys systemu*, wyd. 8, LexisNexis, Warszawa 2005, s. 345; T. Grzegorzczuk, J. Tylman, *Polskie postępowanie karne*, Wydawnictwa Prawnicze PWN, Warszawa 1998, s. 417.

⁹ Por. np. J.L. Ingram, *Criminal evidence*, Elsevier, Amsterdam, Heidelberg, New York, Oxford, Paris, San Diego, San Francisco, Singapore, Sydney, Tokyo 2001, 98–104.

¹⁰ J. Widacki, *Results of polygraph examination: direct Or circumstantial evidence*, *European Polygraph*, 2014, 8, 2 (28), s. 61–67.

opinii biegłego. Analiza praktyki pokazuje, że organy procesowe nie bardzo potrafią poprawnie określić przedmiot i zakres opinii z badań poligraficznych. Zdarzają się więc w postanowieniach zadania dla biegłych sformułowane dziwacznie, często w sposób pozbawiony wszelkiego sensu. Za takie, jak ostatnio wymienione uznać można niewątpliwie na przykład takie jak „zbadanie związku emocjonalnego podejrzanego z zabójstwem”, albo „stwierdzenie, czy w systemie nerwowym badanego są ślady pamięciowe zdarzenia”. Nie wdając się w szczegółowe rozważania, dość zwrócić uwagę, że każdy komu postawiono zarzut popełnienia jakiegoś przestępstwa, bez względu na to, czy jest sprawcą czy nie, ma emocjonalny stosunek do zarzucanego mu czynu i powołującemu biegłego chyba nie o potwierdzenie takiego związku chodzi. Drugi przykład zadania do biegłego jest po prostu śmieszny. System nerwowy człowieka obejmuje cały organizm i poszukiwania w całym organizmie śladów pamięciowych jest chyba przesadą. Równie dobrze można by zapytać, czy „w organizmie badanego znajdują się ślady pamięciowe”. Jak już pytać o ślady pamięciowe wystarczyło by ograniczyć się do samej pamięci badanego. Aby określić, jak powinno się prawidłowo określić przedmiot i zakres ekspertyzy z zakresu badań poligraficznych, wcześniej trzeba przypomnieć, co naprawdę wykazuje badanie poligraficzne.

4. Badanie poligraficzne jest najbardziej popularną instrumentalną metodą detekcji nieszczerości (kłamstwa). W kryminalistyce i psychologii sądowej za kłamstwo uważa się przekazanie fałszywego komunikatu, którego celem jest wytworzenie u odbiorcy komunikatu przekonania, co do którego wysyłający komunikat jest przekonany, że jest fałszywe¹¹. Krótko mówiąc, dla kryminalistyki, a tym samym procesu karnego interesujące jest kłamstwo polegające na tym, że ktoś (osoba podejrzana, podejrzany, oskarżony, świadek) wypowiada świadomie zdanie fałszywe, w intencji, by ktoś (z reguły organ procesowy) uznał je za prawdziwe¹². Kryminalistyka, ani żadna inna dyscyplina naukowa, z psychologia włącznie, nie dysponuje metodami pozwalającymi na bezpośrednie wykrycie kłamstwa jako takiego. Instrumentalne metody detekcji kłamstwa, w tym badanie poligraficzne pozwalają na wykrycie zjawisk towarzyszących kłamstwu, w dodatku takiemu, z którego wykryciem łączą się dla kłamiącego jakieś istotne dla niego konsekwencje. Takimi zjawiskami towarzyszącymi są zmiany emocjonalne, obserwowane za pośrednictwem towarzyszących im zmian fizjologicznych, a także pewien wysiłek intelektualny, który towarzyszy kontrolowaniu się osoby kłamiącej. I to jest wykrywane w badaniu poligraficznym. Wydaje się więc logiczne, że organ procesowy w postanowieniu o powołaniu biegłego, określając zakres ekspertyzy, może pytać o to, co biegły danej specjalności może ustalić. Zatem, skoro w wyniku badań poligraficznych, może ustalić, w oparciu o stwierdzone zmiany emocjonalne, towarzyszące odpowiedziom na pytania krytyczne testu czy badany należy do grupy DI (reagujących tak jak zwykle reagują osoby odpowiadające na te pytania nieszczerze), czy do grupy NDI (reagujących tak jak zwykle reagują osoby odpowiadające na te pytania szczerze), czy do grupy INC (reagujący w sposób nie pozwalający przypisać jednoznacznie do grupy DI lub NDI), to biegły może być pytany w postanowieniu tylko o to, do której grupy zaliczył badanego. Pytanie o „związki emocjonalne”, „obecność w organizmie śladów pamięciowych” pozbawione jest sensu, bowiem biegły z zakresu badań poligraficznych nie to wykrywa czy ustala. Dodać należy, że we współczesnych technikach badań poligraficznych, ta wykryta „nieszczerłość” dotyczy nie odpowiedzi na pojedyncze pytanie testu, ale dotyczy całego obszaru wyznaczonego treścią pytań krytycznych. Nie ustala się zatem, czy na pytanie np. 3 badany odpowiadał szczerze, a na 5 nieszczerze. Takie ustalenia są uprawnione jedynie przy technice klasycznej, dziś już zdecydowanie przestarzałej, nie stosowanej przy badaniach w sprawach karnych, a dopuszczalnej jedynie wyjątkowo w badaniach przed zatrudnieniowych (*pre-employment*) lub kontrolnych (*screening*).

¹¹ Por. np. A. Vrij, *Detecting Lies and Deceit. The Psychology of Lying and Implications for Professional Practice*, John Wiley&Sons, Chichester 2000; por także: J. Widacki, K. Dukała, *Detekcja kłamstwa czyli czego?*, Problemy Kryminalistyki, 2015, 287, s. 6.

¹² Por. J. Widacki, K. Dukała, *op. cit.*, s. 6.

5. Z nieco inną sytuacją mamy do czynienia, jeśli badanie było by wykonane nie techniką pytań kontrolnych¹³, ale techniką Lykkena (w Polsce czasem zwaną „techniką Kulickiego”¹⁴) taką jak Guilty Knowledge Technique (GKT) lub Concealed Information Technique (CIT), które to techniki, nawiasem mówiąc, są nader niepraktyczne, a nadto nie mają wcale wyższej wartości diagnostycznej niż techniki pytań kontrolnych¹⁵. Te techniki badań z założenia nie mają wykrywać nieszczerości (deception), ale „świadomość winnego” (guilty knowledge), lub „ukrywaną informację” (concealed information). Jeśli więc badanie miało by być wykonane jedną z tych technik (GKT lub CIT) to biegłemu można zlecić sprawdzenie, czy u badanego taką „świadomość winnego” lub „fakt ukrywania informacji” wykrył. Przy dużej dozie tolerancji, można by uznać, że ów „związek emocjonalny” jest w jakimś sensie odpowiednikiem „świadomości winnego”, lub „ukrywania posiadanych informacji”, choć było by to wysoce nieprecyzyjne. Ale jeśli badanie wykonane było jedną z technik pytań kontrolnych, to nie ustalano w nim „świadomości winnego” ani „ukrywania informacji”! – tylko szczerść lub nieszczerść badanego w obszarze wyznaczonym pytaniami krytycznymi.

Zatem organ procesowy powołując biegłego z zakresu badań poligraficznych może mu w postanowieniu zlecić ustalenie symptomów szczerści bądź nieszczerści w zakresie objętym pytaniami krytycznymi.

6. Jeśli zatem organ procesowy wydaje postanowienie o powołaniu biegłego w zakresie badań poligraficznych, to upewniwszy się, czy biegły rzeczywiście ma wiadomości specjalne (w pierwszej kolejności, czy dysponuje certyfikatem American Polygraph Association, lub rekomendacją Polskiego Towarzystwa Badań Poligraficznych albo rekomendacją Stowarzyszenia Polskich Poligraferów) powinien – jak tego wymaga art. 194 pkt 2 kpk – poprawnie określić zakres i przedmiot ekspertyzy. Jeśli badanie ma być wykonane techniką pytań kontrolnych (np. najpopularniejszą dziś techniką Utah Zone Comparison Technique) postanowienie powinno w następujący sposób określać przedmiot i zakres ekspertyzy: „**Czy na pytania krytyczne dotyczące (np. zabójstwa Kowalskiego) badany reaguje tak, jak zwykle reagują osoby, które na pytania takie odpowiadają nieszczerze, t.j. kłamią lub zatają fakt posiadania informacji związanych z tym zabójstwem?**”. Jeśli organ procesowy zakłada, że istnieje możliwość wykonania badania techniką GKT lub CIT, to znaczy badany został zatrzymany bezpośrednio przed badaniem, a szczegóły przestępstwa nie były ujawnione w mediach, w związku z czym można zasadnie zakładać, że znajomość szczegółów przestępstwa różnicuje sprawców i świadków od tych, którzy nie mają żadnych wiadomości o przestępstwie można tak sformułować przedmiot ekspertyzy: „**czy badany reaguje tak jak osoba, która zna szczegóły zdarzenia ale temu przeczy?**”. Ponieważ decyzja o tym, którą technikę badania zastosować w konkretnym przypadku, powinna należeć do biegłego, najlepiej, gdy zadanie ekspertyzy (tu: badania poligraficznego) sformułowane jest w sposób wyżej opisany, w wariantcie pierwszym „**czy badany na pytania krytyczne testów dotyczących zabójstwa Kowalskiego reaguje tak, jak osoby, które na pytania krytyczne testów odpowiadają nieszczerze, t.j. kłamią, lub zatają fakt posiadanych informacji związanych z tym zabójstwem?**”. Tak sformułowane zadania biegłego nie przesądzą jaką techniką (CQT czy GKT lub CIT) badanie ma być przeprowadzone.

Po przeprowadzeniu badania i po analizie stwierdzonych reakcji, biegły w opinii powinien napisać odpowiedź na tak postawione pytanie. To znaczy powinien napisać, czy badany reaguje tak jak zwykle reagują osobnicy którzy na pytania krytyczne testów odpowiadają nieszczerze, t.j. kłamią lub zatają fakt posiadania istotnych informacji w tej sprawie.

¹³ Techniki pytań kontrolnych to np. technika Reida, technika Backstera, Federal Zone Comparison Technique, Utah Zone Comparison Technique. Istota tych technik jest stosowanie testów, w których obok pytań krytycznych i obojętnych stosuje się pytania kontrolne. Ocena reakcji polega na porównywaniu reakcji na pytania krytyczne z reakcjami na pytania kontrolne. Współcześnie w świecie, także w Polsce stosuje się przede wszystkim techniki pytań kontrolnych. Najbardziej popularną techniką badania jest technika Utah ZCT.

¹⁴ Rzecz była w Polsce wielokrotnie wyjaśniana i opisywana, ostatnio: J. Widacki, *Spór o technikę badań poligraficznych*, Studia Prawnicze, 2016, 2(19) s. 9–24.

¹⁵ Por. *Meta-analytic Survey of Criterion Accuracy of Validated Techniques*, Polygraph 2011, 20, s. 1.

Biegły powinien też podać podstawy swej konkluzji. Powinien podać jaką technikę badania zastosował, podać liczbową wartość ocenianych reakcji oraz jaką metodę numerycznej oceny zapisów zastosował. Współczesny standard badań poligraficznych wymaga zastosowania numerycznej oceny zapisów. Oceny jakościowe (zwane często eufemistycznie „holistycznymi”) są współcześnie, jako zbyt przestarzałe, a w szczególności zbyt subiektywne niedopuszczalne. Do opinii powinny być dołączone kopie zapisów poligraficznych, tak, by ewentualnie powołani w trybie art. 201 kpk inni biegli mogli ocenić prawidłowość wykonanego badania i prawidłowość ocen reakcji.

7. Wyjaśnienia wymaga, co znaczy użyte w opinii określenie „reaguje tak, jak zwykle reagują osoby, które na pytania krytyczne odpowiadają nieszczerze, to jest kłamią lub zatajają fakt posiadania informacji związanych z przestępstwem, którego dotyczy badanie”? Odnieść się tu trzeba do wartości diagnostycznej badania poligraficznego. Na marginesie warto odnotować, że wartość diagnostyczna badania poligraficznego należy do najlepiej zbadanych w naukach sądowych¹⁶. Wartość ta określana jest jakościowo („jest nie mniejsza niż innych rutynowo stosowanych metod nauk sądowych”¹⁷) lub ilościowo, przez podanie procentów wskazań (diagnoz) trafnych, błędnych i nierozstrzygających¹⁸. Przyjmuje się, że wartość diagnostyczna prawidłowo wykonanych badań poligraficznych, po odrzuceniu wyników nierozstrzygających, wynosi między 85–95% trafnych wskazań (diagnoz). Inaczej mówiąc, szansa błędnego wskazania (diagnozy) wynosi od 5–15%. Jest to porównywalne z wartością diagnostyczną innych rutynowo stosowanych w kryminalistyce i akceptowanych w procesie karnym metod identyfikacji¹⁹. Co więcej, wartość diagnostyczna badania poligraficznego jest wyższa od wartości diagnostycznej ekspertyzy pismo znawczej czy rozpoznania przez świadka w ramach okazania. Trzeba pamiętać, że ta relatywnie wysoka wartość diagnostyczna jest wyliczona dla badań wykonanych którąś z wystandaryzowanych technik. Jeśli badanie wykonane było techniką niestandardową, to jego wartość diagnostyczna jest nieznana. Jeśli więc, po poprawnie wykonanym badaniu poligraficznym, biegły stwierdzi, że osoba badana reagowała na pytania krytyczne tak, jak zwykle reagują osoby, które na te pytania odpowiadają nieszczerze (kłamią, zatajają coś), to znaczy, że na 100 badanych, tak zareagowało by 85–95 kłamiących, a tylko 5–15 szczerych, prawdomównych. I tylko (aż?) tyle może ustalić biegły w badaniu poligraficznym.

8. Dowód z badań poligraficznych podlega zatem, jak każdy dowód ocenie w kontekście innych dowodów, według reguł ustalonych w art.7 kpk. Jest to dowód pośredni (co przyznał m.in. Sąd Najwyższy w orzeczeniu z dnia 29.01.2015 – por. wyżej), a zatem taki, który nie dotyczy bezpośrednio faktu głównego. Ustalenie dokonane w badaniu poligraficznym wymaga dopiero logicznego powiązania z faktem głównym, w szczególności wykazania poprzez inne dowody, że bardziej prawdopodobne jest, że odnotowane reakcje badanego na pytania krytyczne są wynikiem tego, że na pytania krytyczne odpowiadał on nieszczerze, niż że są jakimś artefaktem o nieznanym przyczynie. Może to wbrew pozorom, ale identycznie postępujemy z dowodem z badań DNA. Ujawnienie śladu DNA na odzieży podejrzanego wymaga stwierdzenia, że bardziej prawdopodobne jest, że znalazł się on tam w związku z zabójstwem niż w innych okolicznościach.

¹⁶ Por. J. Widacki, *Wartość diagnostyczna badania poligraficznego i jej znaczenie kryminalistyczne*, Wydawnictwo UJ, Kraków 1977, por. także: J. Widacki, F. Horvath, *An experimental investigation of the relative validity and utility of the polygraph technique and Three others common methods of criminal identification*, Journal of Forensic Sciences, 1978, 23, 3, s. 596–601; a ostatnio: *American Polygraph Association Report*, Polygraph 2011 (special issue).

¹⁷ Por. np. J. Widacki, *op. cit.*; J. Widacki, F. Horvath, *op. cit.*

¹⁸ Tak na przykład uczyniono we wspomnianym wyżej Raporcie APA.

¹⁹ Por. J. Widacki, F. Horvath, *op. cit.*

Anna Szuba-Boroń, Marcin Gołaszewski

Badania poligraficzne funkcjonariuszy państwowych w Polsce*

Pierwszą polską służbą mundurową, która weszła w posiadanie poligrafu, była Wojskowa Służba Wewnętrzna (WSW) – formacja łącząca w sobie funkcje policji wojskowej i kontrwywiadu wojskowego. Pod koniec lat 60. XX wieku zakupiła poligraf Keelera (model 6308). Nieco później poligraf nabyło Ministerstwo Spraw Wewnętrznych, w strukturze którego były: cywilny wywiad (Departament I) i kontrwywiad (Departament II).

W obydwu instytucjach, poligrafy służyły przede wszystkim do celów szkoleniowych funkcjonariuszy, a zapewne także agentów wywiadu zapoznawano z takimi badaniami, zakładając, że mogą oni być w przyszłości poddani takim badaniom przez służby amerykańskie. W WSW wykorzystywano poligraf w sprawach karnych, w śledztwach prowadzonych przez prokuratury wojskowe, w szczególności w sprawach kradzieży broni. Eksperci WSW byli też powoływani przez prokuratury powszechne (nie wojskowe) do wykonywania badań poligraficznych w różnych najpoważniejszych sprawach karnych, w tym przede wszystkim w sprawach o zabójstwa.

Do roku 1989 Milicja Obywatelska (Policja) nie dysponowała żadną komórką badań poligraficznych, korzystając z usług ekspertów WSW lub ekspertów uniwersyteckich. Nie wykonywano też w tym czasie badań poligraficznych kandydatów do służb policyjnych ani specjalnych.

Zasadnicza zmiana nastąpiła po zmianie ustroju i upadku komunizmu po roku 1989. Uniezależnienie od ZSRR, reforma służb specjalnych i policyjnych oraz nawiązanie nieskrępowanej współpracy w różnych dziedzinach z państwami zachodnimi, w tym Stanami Zjednoczonymi pozwoliło na szersze niż dotąd stosowanie badań poligraficznych i wykorzystanie wzorów zachodnich.

Policja, która zastąpiła Milicję Obywatelską zakupiła pierwsze poligrafy i zorganizowała zarówno centralny ośrodek badań poligraficznych w Warszawie, jak i 4 ośrodki terenowe. Również utworzony w 1990 roku Urząd Ochrony Państwa uruchomił swoją komórkę badań poligraficznych. Takie komórki powstały także w Straży Granicznej i Żandarmerii Wojskowej (Policji Wojskowej). Po reorganizacji służb specjalnych: cywilnych – Urząd Ochrony Państwa podzielono na Agencję Bezpieczeństwa Wewnętrznego i Agencję Wywiadu¹, utworzono Centralne Biuro Antykorupcyjne (CBA)², a w miejsce Wojskowych Służb Informacyjnych (WSI) utworzono Służbę Kontrwywiadu Wojskowego i Służbę Wywiadu Wojskowego (SKW i SWW)³.

Wszystkie te służby, tak policyjne jak i specjalne, dysponują swoimi komórkami badań poligraficznych. Wszystkie wykorzystują badania tak w sprawach śledczych jak i operacyjnych, na swój wewnętrzny użytek.

* Artykuł ukazał się w języku rosyjskim w: *Primienienije poligrafa w bywszem wostocznom blokcie*, red. J. Wiśniewski, W. Szapowałow, I. Usikow, Oficyna Wydawnicza AFM, Kraków 2017, s. 145–154.

¹ Ustawa z dnia 24 maja 2002 r. o Agencji Bezpieczeństwa Wewnętrznego oraz Agencji Wywiadu (Dz.U. z 2002 r. nr 74, poz. 676).

² Ustawa z dnia 9 czerwca 2006 r. o Centralnym Biurze Antykorupcyjnym (Dz.U. z 2006 r. nr 104, poz. 708).

³ Ustawa z dnia 9 czerwca 2006 r. o Służbie Kontrwywiadu Wojskowego oraz Służbie Wywiadu Wojskowego (Dz.U. z 2006r. nr 104 poz. 709).

Badania poligraficzne znajdują zastosowanie w postępowaniach kwalifikacyjnych i kontrolnych (screening) większości służb policyjnych i specjalnych. Procedury przyjęć (rekrutacyjne) do służby są określone ustawowo i przewidują obowiązkowe badania poligraficzne kandydatów. Wykorzystuje się je również do badania lojalności funkcjonariuszy służb oraz żołnierzy. Liczba badań na użytek wewnętrznych potrzeb służb policyjnych i specjalnych nie jest znana, można szacować, że jest ich co najmniej kilkakrotnie więcej niż badań wykonywanych w sprawach karnych.

Poligraficznym badaniom przedzatrudnieniowym poddawani są kandydaci do służby w: Policji, Centralnym Biurze Antykorupcyjnym, Agencji Wywiadu, Agencji Bezpieczeństwa Wewnętrznego, Straży Granicznej, Żandarmerii Wojskowej, Służbie Wywiadu Wojskowego, Służbie Kontrwywiadu Wojskowego, Krajowej Administracji Skarbowej, a także funkcjonariusze i pracownicy tych instytucji.

Celem badań przedzatrudnieniowych jest sprawdzenie przydatności danego funkcjonariusza do służby na danym stanowisku lub w określonej komórce organizacyjnej i eliminacja tych spośród kandydatów, którzy z określonych powodów nie powinni pełnić służby lub pracować na określonym stanowisku.

W przypadku ABW i AW, SKW i SWW⁴, CBA oraz Straży Granicznej⁵ badanie poligraficzne może być przeprowadzone w przypadku kandydatów na stanowiska, które wymagają szczególnych umiejętności lub predyspozycji (de facto są to niemal wszystkie stanowiska).

W służbach specjalnych o skierowaniu funkcjonariusza na badanie poligraficzne decydują szefowie tych służb, którzy nie muszą uzasadniać swojej decyzji. Warto przy tym nadmienić, że ewentualna odmowa poddania się badaniu byłaby potraktowana jako niewykonanie polecenia służbowego, czyli delikt dyscyplinarny. Z kolei w przypadku Straży Granicznej – przeprowadzenie badania poligraficznego wobec funkcjonariusza możliwe jest w związku ze skierowaniem na szkolenie oraz postępowaniem w sprawie o przestępstwo popełnione w związku z wykonywaniem czynności służbowych⁶.

W Krajowej Administracji Skarbowej badaniom poligraficznym poddawani są kandydaci przewidziani do służby w większości komórek organizacyjnych Służby Celno-Skarbowej⁷. W celu sprawdzenia przydatności do służby na danym stanowisku badaniom mogą być poddani także funkcjonariusze (chodzi przede wszystkim o tych, którzy są uprawnieni do wykonywania czynności operacyjno-rozpoznawczych). Warto przy tym zwrócić uwagę na kryteria, które powinna spełnić osoba przeprowadzająca badania poligraficzne w KAS. Zgodnie z rozporządzeniem Ministra Rozwoju i Finansów – poligrafer „posiada tytuł magistra, ukończył kurs specjalistyczny i uzyskał stosowny certyfikat autoryzowany lub uznany przez American Polygraph Association (APA) lub kurs z zakresu badań poligraficznych na poziomie co najmniej podstawowym, lub posiada odpowiednie, nabyte w innej drodze umiejętności prawidłowego przeprowadzania badania poligraficznego i interpretacji jego wyników poświadczone przez polskie organizacje z zakresu badań poligraficznych, w tym Polskie Towarzystwo Badań Poligraficznych”⁸. Wymienione rozporządzenie jest najnowszym dokumentem tej rangi w Polsce, który reguluje kwestie kwalifikacji poligrafera.

Przy zatrudnianiu funkcjonariuszy Policji badania poligraficzne mają służyć określeniu predyspozycji danego funkcjonariusza do pełnienia służby na określonych stanowiskach lub w określonych

⁴ Rozporządzenie Ministra Obrony Narodowej z dnia 26 lipca 2006 r. w sprawie postępowania kwalifikacyjnego wobec kandydatów do służby w Służbie Kontrwywiadu Wojskowego (Dz.U. z 2006 r., poz. 978). Rozporządzenie Ministra Obrony Narodowej z dnia 26 lipca 2006 r. w sprawie postępowania kwalifikacyjnego wobec kandydatów do służby w Służbie Wywiadu Wojskowego (Dz.U. z 2006 r., poz. 979).

⁵ Ustawa z dnia 12 października 1990r. o Straży Granicznej (Dz.U. z 1990 r. nr 78, poz. 462).

⁶ Zgodnie z art. 1 ust. 2a. Ustawy o Straży Granicznej – Straż Graniczna prowadzi postępowania w sprawach rozpoznawania, zapobiegania i wykrywania przestępstw określonych w art.: 228 (łapownictwo bierne), 229 (łapownictwo czynne) i 231 (nadużycie uprawnień, niedopełnienie obowiązków) Kodeksu karnego, popełnionych przez funkcjonariuszy Straży Granicznej w związku z wykonywaniem czynności służbowych. (W takim przypadku na badanie kieruje funkcjonariusza Komendant Główny Straży Granicznej).

⁷ Ustawa z dnia 16 listopada 2016 r. o Krajowej Administracji Skarbowej (Dz.U. z 2016 r., poz. 1947).

⁸ § 7 Rozporządzenia Ministra Rozwoju i Finansów z dnia 10 kwietnia 2017r. w sprawie przeprowadzania badania psychofizjologicznego, testu sprawności fizycznej oraz badania psychologicznego funkcjonariuszy Służby Celno-Skarbowej (Dz.U. z 2017 r., poz. 805).

komórkach organizacyjnych, w szczególności: ustaleniu lojalności wobec służby, czerpania nieuprawnionych korzyści w dotychczasowej służbie oraz ustaleniu patologii i niepożądanych uzależnień w służbie. Podczas badania poligraficznego nie zadaje się pytań dotyczących wyznania oraz preferencji seksualnych i politycznych. Przebieg badania jest rejestrowany (obraz i dźwięk), a jeśli wynik badania jest niejednoznaczny można takie badanie powtórzyć tylko jeden raz, w terminie 30 dni⁹.

Badania kontrolne mają głównie za cel sprawdzenie, czy funkcjonariusze od czasu ostatniego badania nie ujawnili informacji niejawnych; czy nie nawiązali niepożądanych kontaktów; czy nie popełnili przestępstw lub przewinień dyscyplinarnych itp.

Aktualnie – spośród wszystkich służb wykonujących zadania w zakresie bezpieczeństwa państwa – jedynie w przypadku Biura Ochrony Rządu brak regulacji ustawowych, które wprowadzałyby obowiązki i w ogóle – możliwość – badań poligraficznych wobec kandydatów do służby i funkcjonariuszy. Ministerstwo Spraw Wewnętrznych i Administracji zapowiada jednak reformę, w ramach której planowane są m.in. nowe zasady rekrutacji do służby, obejmujące także badanie poligraficzne.

Nie ulega wątpliwości, że skoro uprawnienie do badań poligraficznych kandydatów do służby i funkcjonariuszy ma oparcie w przepisach rangi ustawowej – to znaczy, że żaden inny urzędnik państwowi nie mogą być poddawani takim badaniom, jeśli ustawa tego nie przewiduje.

Służby policyjne i specjalne, które realizują czynności operacyjno-rozpoznawcze, mogą ponadto wykonywać badania poligraficzne wobec osób udzielających pomocy przy takich czynnościach.

Badania poligraficzne są dopuszczalne także w postępowaniach dyscyplinarnych prokuratorów. Przepisy ustawy Prawo o prokuraturze¹⁰ (w rozdziale, który dotyczy odpowiedzialności karnej, dyscyplinarnej i służbowej prokuratorów), mówią o tym, że rzecznik dyscyplinarny w trakcie prowadzonego postępowania dyscyplinarnego może w celu ograniczenia kręgu osób podejrzanych o popełnienie przewinienia dyscyplinarnego zawierającego znamiona przestępstwa ujawnienia informacji z postępowania karnego stanowiących informacje niejawne o klauzuli tajności „tajne” lub „ściśle tajne”, powołać biegłego w celu zastosowania wobec prokuratora posiadającego dostęp do tych informacji, za jego zgodą, środków technicznych mających na celu kontrolę nieświadomych reakcji organizmu¹¹.

Wieloletnie doświadczenia pozostałych polskich służb mundurowych dowodzą wysokiej wartości badań psychofizjologicznych z wykorzystaniem poligrafu, ponieważ dzięki nim pozyskuje się cenne informacje, trudne lub wręcz niemożliwe do uzyskania w ramach innych dostępnych procedur. Mimo to, badania przeprowadzane w instytucjach państwowych wywołują niekiedy rozmaite kontrowersje, które występują na dwóch płaszczyznach. Pierwsza związana jest z dopuszczalnością takich badań w sensie formalno-prawnym i etycznym, a wątpliwości w tym zakresie zgłaszane są przez osoby spoza służb, z naciskiem na rzekomo wysoki poziom błędów badań poligraficznych. Natomiast druga płaszczyzna kontrowersji obejmuje zagadnienia standaryzacji badań i politykę zapewnienia jakości. Ta część dyskusji toczy się międzyresortowo i w gronie funkcjonariuszy poszczególnych instytucji.

Kwestia dopuszczalności badań poligraficznych kandydatów do służby i funkcjonariuszy była już w Polsce przedmiotem badań Trybunału Konstytucyjnego¹², który nie dopatrywał się niekonstytucyjności przepisów ustawy pozwalających na takie badania w Policji. Organom państwowym wolno czynić tylko to, co jest przez ustawę nakazane lub dozwolone. Zakres tych kompetencji został wyżej wskazany dla poszczególnych instytucji i nie powinien budzić wątpliwości.

⁹ Rozporządzenie Ministra Spraw Wewnętrznych i Administracji, z dnia 20 marca 2007r. w sprawie trybu i warunków ustalania zdolności fizycznej i psychicznej policjantów do służby na określonych stanowiskach lub w określonych komórkach organizacyjnych jednostek Policji (Dz.U. z 2007 r., nr 62, poz. 423).

¹⁰ Ustawa z dnia 28 stycznia 2016 r. Prawo o prokuraturze (Dz.U. 2016, poz. 177).

¹¹ *Ibidem*, art. 154.

¹² Wyrok Trybunału Konstytucyjnego z dnia 23 lutego 2010r. K. 1/2008 LexPolonica nr 2144497, OTK ZU 2010/2A, poz. 14.

Oponenci często stawiają przed badaniami poligraficznymi nierealny postulat bezwzględnej trafności, przytaczając jednocześnie zaniżone dane (np. 75%)¹³. Tymczasem należy podkreślić, że badania poligraficzne mają relatywnie bardzo wysoką i prawdopodobnie najlepiej zbadaną wartość diagnostyczną względem innych metod wykorzystywanych w kryminalistyce. Potwierdzona badaniami naukowymi średnia dokładność wszystkich typów testów poligraficznych wynosi 87,1%¹⁴. Dla testów diagnostycznych (jednoproblemowych) jest to 92,1%, zaś dla przesiewowych (wieloproblemowych) – 85%. Dane dotyczące dokładności technik badań poligraficznych wynikają z rozmaitych badań naukowych prowadzonych na przestrzeni lat w XX i XXI wieku. Brane są pod uwagę te, które spełniają naukowe standardy i zostały opublikowane w niezależnie recenzowanych czasopiśmie i monografiach naukowych. W rezultacie metaanalizy danych pochodzących z wielu badań powstały m.in. dwa ważne raporty podsumowujące średnią dokładność badań poligraficznych ogółem, a także różnych typów badań i poszczególnych technik badawczych. Jeden ze wspomnianych raportów wydała amerykańska National Research Council w 2003 roku. Potwierdzono wówczas wskaźnik dokładności badań poligraficznych prowadzonych w „konkretnych sprawach” – wynoszący ok. 90%. Drugi raport został opracowany przez American Polygraph Association w 2011 r. i ten uznaje się aktualnie za najlepszy punkt odniesienia. Jak w każdej innej dziedzinie nauki, dane opublikowane w raporcie należy uznać za obowiązujące, dopóki ewentualne nowe wyniki dowodów naukowych nie zaowocują istotną zmianą.

Znacznie bardziej zasadny w kontekście dokładności testów stosowanych w badaniach poligraficznych jest postulat minimalizacji ryzyka błędu, wyrażony w marcu 2017r. m.in. przez przedstawiciela Helsińskiej Fundacji Praw Człowieka¹⁵. Temu postulatowi wychodzą naprzeciw standardy profesjonalne¹⁶. Jeśli jakiś test, a tym bardziej całe badanie poligraficzne, jest na pewnym etapie nierozstrzygnięte, należy wykonać dodatkowe testy (w ramach tego samego badania lub kolejnego – powtórzonego). Z kolei tzw. podejście „sukcesywnego pokonywania przeszkód” wymaga, aby przy niekorzystnym dla badanego wyniku testu przesiewowego przeprowadzić test diagnostyczny w celu weryfikacji powstałych wątpliwości. Technicznie oznacza to, że po teście o relatywnie mniejszej dokładności (średnio 85%), powinien nastąpić test bardziej dokładny (średnio 92,1%).

Należy też wyraźnie podkreślić, że nie wolno stawiać znaku równości między stanowiskami w instytucjach rządowych i organach egzekwowania prawa a miejscami pracy w zwykłych urzędach czy zakładach pracy. Nikomu nie trzeba chyba tłumaczyć, że z punktu widzenia bezpieczeństwa państwa, lepiej odrzucić pewien procent kandydatów – być może nie do końca słusznie ocenionych – niż przyzwolić na przemknienie przez rekrutacyjne sito jednostce, która może wyrządzić poważne szkody Rzeczypospolitej Polskiej. Interes związany z bezpieczeństwem państwa i ochroną informacji niejawnych ma pierwszeństwo przed innymi prawnie chronionymi interesami¹⁷.

Zasady przeprowadzania badań poligraficznych określają ogólnodostępne standardy profesjonalnych i normalizacyjnych organizacji o zasięgu międzynarodowym (American Polygraph Association, ASTM International) oraz krajowym (w tym: Polskie Towarzystwo Badań Poligraficznych). Ponadto istnieją wewnętrzne zarządzenia i instrukcje w ramach różnych instytucji państwowych, w których przeprowadza się badania poligraficzne. Standardy zostały omówione w polskojęzycznych

¹³ M.in. P. Słowik, *Rekrutacja do Straży Granicznej: Wariograf wątpliwy, ale nie dla MSWiA*, „Gazeta Prawna”, <http://serwis.gazetaprawna.pl/praca-i-kariera/artykuly/1026987,rekrutacja-do-strazy-granicznej-badanie-wariografem-w-sadzie.html> [dostęp: 26.03.2017].

¹⁴ American Polygraph Association, *Meta-Analytic Survey of Criterion Accuracy of Validated Techniques*, „Polygraph”, t. 40, nr 4, 2011.

¹⁵ P. Słowik, *Rekrutacja do Straży Granicznej...*

¹⁶ Szerzej o standardach badań poligraficznych w: M. Gołaszewski (red.), *Współczesne standardy badań poligraficznych*, Warszawa 2013; J. Widacki (red.), *Badania poligraficzne w Polsce*, Kraków 2014.

¹⁷ Zob. art. 24. ust. 4 Ustawy z dnia 5 sierpnia 2010r. o ochronie informacji niejawnych (Dz.U. z 2010 r., nr 182, poz. 1228).

publikacjach naukowych¹⁸, więc są powszechnie dostępne i bariera językowa nie jest dostatecznym usprawiedliwieniem, aby je ignorować.

Do badań kadrowych typu *screening* rekomenduje się te techniki, które są potwierdzone naukowo i mają poziom dokładności istotnie wyższy niż statystyczna szansa. Mogą być stosowane wraz z podejściem tzw. sukcesywnego pokonywania przeszkód.

Jak wiadomo – poszczególne testy, w powiązaniu z określonym systemem analizy danych testowych, różnią się między sobą nie tylko średnią dokładnością, ale także czułością i swoistością. Optymalnym rozwiązaniem wydaje się dobór takiej techniki, która ma te parametry zrównoważone. Polityka danej instytucji może jednak przewidywać, że celem nadrzędnym jest jak najwyższa czułość na kłamstwo, nawet kosztem zdolności testu do potwierdzenia prawdomówności badanego. Regulacji mogą podlegać także progi decyzyjne przy ocenach numerycznych – w zależności od przyjmowanej tolerancji błędu. Standardowo przyjmuje się na przykład tolerancję błędu przy wynikach korzystnych dla badanego na poziomie 0,1. System można by jednak uszczelnić, obniżając tę tolerancję do 0,05 (tak jak to jest w przypadku wyników wskazujących na nieszczerść).

Możliwość wyboru spośród kilkunastu różnych testów nie powinna jednak oznaczać zupełnej dowolności. Na uwagę zasługuje tu w szczególności kwestia stosowania testu R/I (tzw. klasycznego), który powoli odchodzi do historii, a relatywnie niewysoka dokładność (na poziomie ok. 75%), w połączeniu z zatrważająco niską swoistością i wyłącznie jakościową (globalną) metodą ewaluacji wykresów, poddają w wątpliwość sens jego stosowania w ogóle. W pewnych warunkach bywa użyteczny, jednak w normalnej sytuacji trudno zrozumieć jakiekolwiek argumenty za wyborem tego testu (poza prostotą techniki i wygodą badającego) – skoro istnieją inne, o korzystniejszych parametrach pod każdym względem. Podobne zastrzeżenia może budzić także kwestia stosowania testów szczytowego napięcia (POT) w taki sposób, że ich wyniki stanowią podstawę wydania opinii o wiarygodności osoby badanej. W tym miejscu warto więc przypomnieć, że test typu POT w wersji z poszukiwanym rozwiązaniem ma charakter wyłącznie pomocniczy.

Na gruncie tego typu wątpliwości pojawia się temat tzw. kontroli jakości czy – delikatniej rzecz ujmując – polityki zapewnienia jakości. W Stanach Zjednoczonych takie procedury z powodzeniem funkcjonują. Pomagają eliminować patologie (np. wyniki na zamówienie), ograniczać subiektywizm badających, mobilizować do doskonalenia zawodowego, a z drugiej strony zapewniają też samym badającym ochronę przed naciskami zewnętrznymi i próbami podważania opinii. W Polsce, w różnych instytucjach i zróżnicowanym zakresie, wprowadzono wewnętrzne procedury, obejmujące między innymi tzw. ślepą ocenę poligramów, sprawdzania wybranych badań w całości („na żywo” lub z odtworzenia) i nadzór merytoryczny właściwych przełożonych. Brakuje natomiast organu kontrolnego, który koordynowałby programy badań poligraficznych na poziomie centralnym, międzyresortowym i odpowiadałby równocześnie za proces szkolenia ekspertów i badania naukowe – na wzór National Center for Credibility Assessment w Stanach Zjednoczonych. Z podobnym problemem spotykamy się również przy badaniach w procesie karnym, gdzie wobec braku procedur licencyjnych, działalność biegłych pozostaje jedynie pod formalną kontrolą sądową i ewentualnie – naukową oraz środowiskową, ale część biegłych pozostaje poza strukturami stowarzyszeń zawodowych i nie jest zainteresowana utrzymywaniem jakichkolwiek kontaktów z ekspertami.

Niezależnie od wskazanych kontrowersji i problemów można podsumować, że badania poligraficzne w polskich instytucjach państwowych mają już szeroko rozwiniętą praktykę. Wydaje się też, że poligraferzy zapracowali na ugruntowaną pozycję i renomę w odbiorze osób kierujących służbami i tzw. resortami siłowymi.

¹⁸ M. Gołaszewski (red.), *Współczesne standardy badań poligraficznych*, Warszawa 2013, „Biblioteka Przeglądu Bezpieczeństwa Wewnętrznego”, nr 4; M. Gołaszewski, M. Widacki, *Aktualny standard badań poligraficznych a praktyka polska*, w: J. Widacki (red.), *Badania poligraficzne w Polsce*, Kraków 2014.

Podsumowanie

W ramach projektu udało nam się uporządkować terminologię. Określić, co rozumiemy przez „detekcję kłamstwa”, uzasadnić, dlaczego powinno się używać terminu „poligraf”, a nie „wariograf” – konsekwentnie tej nazwy używaliśmy w naszych badaniach i opisie ich wyników. Określiśmy zakres subiektywizmu w badaniach poligraficznych i podaliśmy sposoby jego określenia i mierzenia. Wybraliśmy technikę badania poligraficznego (technikę pytań kontrolnych, w odmianie Utah Zone Comparison Technique) i uzasadniliśmy, dlaczego nie używaliśmy techniki Lykkena. W tym zakresie otwarto drogę do dalszych badań problemu subiektywizmu w ekspertyzie poligraficznej, które w oparciu o zgromadzony materiał eksperymentalny, można będzie kontynuować w ramach pracy doktorskiej jednego z uczestników projektu (Marcina Gołaszewskiego). Przeprowadzony eksperyment dostarczył nam materiału do dalszych badań. Przeanalizowaliśmy aktualny stan wiedzy dotyczący nieinstrumentalnych metod detekcji kłamstwa, porównując możliwości tych technik z możliwościami badań poligraficznych. Bardzo ważną częścią naszych badań była próba wykorzystania do detekcji kłamstwa obserwacji zmian temperatury twarzy, rejestrowanych kamerą termowizyjną. Opracowaliśmy własną, nową metodę przekształcania obrazu kamery termowizyjnej w wykres, który można porównywać z zapisem reakcji rejestrowanych przez klasyczny poligraf – szczególnie reakcji elektrodermalnej (*galvanic skin response* – GSR, *electrodermal activity* – EDA). Wyniki naszego eksperymentu w tym zakresie mogliśmy porównać z wynikami badań eksperymentalnych autorów zagranicznych i ocenić stopień zaawansowania tej metody oraz możliwość wykorzystania jej w praktyce. Opracowaliśmy metodykę, którą można będzie doskonalić w dalszych badaniach. Rozważyliśmy, na ile dotychczasowa reglamentacja prawna i etyczna jest wystarczająca dla przyszłych możliwości technicznych detekcji kłamstwa i zasugerowaliśmy konieczność wprowadzenia nowych regulacji i ograniczeń, motywowanych ochroną podstawowych praw i wolności ludzkich.

